PROTEINS AS BULDING BLOCKS FOR BIOLOGICAL NANOMATERIALS

A thesis submitted in partial fulfilment of the requirements for the

Degree of Doctor of Philosophy in Biochemistry

2016

Helen Ashmead
University of Canterbury

Acknowledgements

I would first like to thank Juliet Gerrard. I could not have hoped for a better supervisor and mentor. You have provided me with so many wonderful opportunities throughout my thesis and I will always be grateful for your encouragement and belief in me. I would also like to thank my co-supervisor, Leonardo Negron, who helped me through the first stages of my thesis and made me feel so welcome when I first arrived in Wellington. To Grant Pearce, my supervisor at UC, thank you for your patience and guidance throughout and for all of the AUC analysis you helped me to complete. To Celine Valery and Vic Arcus, thank you for your supervision during the first half on this thesis.

Thank you to Jackie Healey for producing so many wonderful TEM images and for always being there for help and advice when I visited the lab at UC. To the other students at UC, whose lively discussions (and delicious cake) made my time there so enjoyable. I would especially like to thank Amy Yewdall, Amy Phillips and Rishi Pande. Thank you for your expertise, advice and, for making me feel so welcome. Thank you to Ren Dobson and all the other academic staff at BIC for all your help.

Thank you to all of the staff at IBT, Callaghan Innovation, who helped me with technical support and guidance, especially Shuguang Zhang, who was always willing to give me assistance and who helped in the design of the plasmids. I would also like to thank my fellow student at IBT, Sam Kim, who was always there to talk and who taught me so much about molecular biology.

Completing this thesis would not have been possible without my friends outside of university. To my flatmates, who made Wellington a home away from home and who have become my surrogate family in NZ. And a special thanks to Rob and Kate, who supported me through the writing of my thesis and helped me to proofread and format in the last stages. I will forever be grateful for your help getting me over the line.

Finally, I would like to thank my wonderful family for their unfailing belief in me. The decision to move to the other side of the world to complete my studies was made all the more difficult knowing that I would be leaving you. You have always been there for me throughout my time in NZ, thank you for letting me know you are thinking of me and sending daily messages of love and support.

Table of contents

Abbreviations

1D	1 dimensional
2D	2 dimensional
3D	3 dimensional
Å	angstrom
A280	absorbance at 280 nm
AFM	Atomic force microscopy
AT	adenine thymine
ATP	Adenosine Triphosphate
AU	Absorbance units
AUC	Analytical ultracentrifugation
Au NPs	gold nanoparticles
BIC	Biomolecular Interaction Centre
BLAST	Basic local alignment search tool
bp	base pair
BSA	bovine serum albumin
B-type	β interface
*Bt*Prx3	bovine peroxiredoxin 3
°C	degree Celsius
CD	circular dichroism
CHCA	α-cyano-4-hydroxycinnamic acid
CV	column volume
C_p	peroxidatic cysteine
CPN	cyclic peptide nanotubes
C_r	resolving cysteine
Cryo-EM	cryo-electron microscopy
Cys-S_pOH	peroxidatic cysteine sulfenic acid
Cys-S_pO_2H	peroxidatic cysteine sulfenic acid
d	derivative
Da	Daltons
D_H	hydrodynamic radius

dH2O	distilled water
DLS	dynamic light scattering
Dmax	maximal particle diameter
DMSO	dimethyl sulfoxide
dn/dc	refractive index increment
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
ε280	extinction coefficient
FF	fully folded
FPLC	fast protein liquid chromatography
g	grams
GA	glutaraldehyde
GST	glutathione S-transferase
GTE	glucose-Tris-EDTA
h	hour
H_2O	water
HBP23	heme-binding protein 23
HCl	hydrochloric acid
HEPES	N-2-hydroxyethlpiperazine-N'-2-ethane sulphonic acid
HPLC	high pressure liquid chromatography
H-NS	histone-like nucleoid structuring protein
HMW	high molecular weight
HRP	horse radish peroxides
*Hs*Prx3	human peroxiredoxin 3
*Hs*Prx3-6his	His_6 tagged human peroxiredoxin 3
*Hs*Prx1	human peroxiredoxin 1
IMAC	Immobilised metal affinity chromatography
IPA	Isopropyl alcohol
IPTG	isopropyl β-D-1-thiogalactopyranoside
IHF	integration host factor

K_{av}	size exclusion distribution coefficient
kDa	kiloDaltons
K-turn	kink turn motif
L	litre
LA	lysogeny broth agar
LB	lysogeny broth
*Lb*Prx1m	*Leishmania braziliensis* Prx 1
LCMS	liquid chromatography mass spectroscopy
LMW	low molecular weight
Lsr2	Protein encoded by the actinobacterial *lsr2* gene
LU	locally unfolded
Lys	lysine
µg	micrograms
µL	microlitre
µM	micromolar
MAK	Methionine Alanine Lysine
MALDI-TOF	matrix assisted laser desorption/ionization time of flight
MBP	maltose binding protein
mL	millilitre
mM	millimolar
min	minute
mTXNPx	mitochondrial tryparedoxin-peroxidase
MMW	medium molecular weight
MOPS	3-morpholinopropanesulfonic acid
MS	mass spectroscopy
M. smegmatis	*Mycobacterium smegmatis*
M. tuberculosis	*Mycobacterium tuberculosis*
MW	molecular weight
MWCO	molecular weight cut-off
MWL	molecular weight ladder
NAP	nucleoid associated protein
nm	nanometres
NPs	nanoparticles

NMR	Nuclear magnetic resonance
NR	non-reduced
Nterm-Lsr2	N-terminal domain of Lsr2
OD	optical density
O/N	overnight
PdI	Polydispersity index
%	percent
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffer saline
PDB	protein data bank
PEG	Polyethylene glycol
*Pf*AOP	*Plasmodium falciparum* antioxidant protein
pI	isoelectric point
*Py*Prx1a	*Plasmoddium yoelii* Prx 1
*Py*Prx1a^{N*}	N-terminal histidine tagged *Plasmoddium yoelii* Prx 1
ProtParam	Protein parameters
Prx	peroxiredoxin
QCM-D	Quartz crystal microbalance with dissipation
Red	Reduced
RFU	Relative fluorescence unit
Rg	radius of gyration
RMSD	root mean square deviation
RNA	ribonucleic acid
RNP	ribonucleoprotein
ROS	reactive oxygenated species
R-type	Ring interface
ROX	reactive oxygen species
rpm	revolutions per minute
S	Svendberg sedimentation coefficient
s	second
S200	Superdex 200 matrix
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography

*Sm*Prx1	*Schistosoma mansoni* Prx1
*sj*GST	*Schistoma japonicum* gluthanione S-transferase
Srx	sulfiredoxin
STEM	scanning tunnelling electron microscopy
SWCNT	single walled carbon nanotubes
TAE	Tris-Acetate EDTA
TCEP	tris(2-carboxyethyl)phosphine
TE	Tris-HCl-EDTA
TEM	transmission electron microscopy
TEV	tobacco etch virus
Tpx	thiol peroxidase
TRAP	tRNA attenuation protein
Trx	thioredoxin
Tris	tris(hydroxymethyl)aminoethane
U	unit
uH2O	ultra pure water
v/v	unit volume per unit volume
Vo	void volume
WT	wild type
w/v	unit weight per unit volume
w/w	unit weight per unit weight
Z-ave	Z average
ZYM-5052	autoinduction media

This thesis had contributed to the following publications

Ashmead HM, Negron L, Webster K, Arcus V, Gerrard JA: Proteins as supramolecular building blocks: Nterm-Lsr2 as a new protein tecton. *Biopolymers* 2015, 103(5):260-270.

Domigan LM, Ashmead HM, Dimartino S, Malmstrom J, Pearce GF, Blunt M, Williams DE, Gerrard JA: Triggered formation of supramolecular protein structures on gold surfaces. *In preparation*

Abstract

Within the field of nanoscience there is a growing interest in the use of biological molecules such as proteins, peptides and nucleic acids, to create materials with complex structure and function. These molecules self-assemble to form a huge variety of functional nanoscale structures in biological systems. Through the work of supramolecular chemistry there is now a greater understanding of how to control the assembly of molecules into functional entities. The aim of this thesis is to combine the fields of biology and supramolecular chemistry to increase the repertoire of available protein building blocks for use in bionanotechnology. Three proteins have been selected due to their ability to self assemble in nature, Lsr2, the N-terminal domain of Lsr2 (Nterm-Lsr2) and human peroxiredoxin 3 (*Hs*Prx3).

Lsr2 is a small DNA binding protein that has the innate ability to self-assemble *in vivo* via a number of routes. This makes it an ideal candidate as a biological building block for forming nanomaterials *in vitro*. The truncated version of Lsr2 (Nterm-Lsr2) oligomerises into linear chains via anti-parallel β-sheet formation between the extended N-termini of neighbouring dimers. This process was facilitated *in vitro* using low concentrations of trypsin which removed three N-terminal residues, Met^1, Ala^2 and Lys^3, allowing an inter-dimer anti-parallel β-sheet to form. Using trypsin to initiate assembly led to unwanted proteolysis at additional lysine and arginine residues within the polypeptide chain. A novel and more controlled method of initiating assembly was developed by replacing three N-terminal residues with an enteropeptidase recognition site. Enteropeptidase cleaves specifically at Lys^4 leaving an identical sequence to the native protein when treated with trypsin but without additional fractionation. This allows the formation of an ordered network of "spaghetti-like" fibres. The structures alternate between a tetramer and high molecular weight oligomers in response to variations in pH. Further control is exerted over the system by exploiting the open symmetry of the assemblies. Increasing and decreasing the protein tecton concentration leads to the formation of larger and smaller structures respectively.

Wild type *Hs*Prx3 has been shown to assemble into one dimensional tubes at acidic pH (pH 4.2). By incorporating the N-terminal histidine tag and linker sequence (*Hs*Prx3-6his) into the current assembly system, a number of novel oligomerisation routes were developed. The presence of the tag and linker stabilises the dodecameric toroid leaving an ideal tecton from

which higher ordered structures can form. The increased pH sensitivity of *Hs*Prx3-6his allows the formation of tubes at pH 7.4. The size of the assembly was further controlled by small changes in pH, with the tube length increasing with decreasing pH values. It is proposed that electrostatic interactions at the ring interface are driving the assembly. Increasing the salt concentration, thereby disrupting these electrostatic interactions, caused the tubes to dissociate into single rings. Increasing the histidine content within the tag led to the formation of longer tubes, suggesting that the presence of the non-native histidine residues is the origin of the pH sensitivity. The metal-binding imidazole side groups of the histidine tag were also utilised to stabilise stacks of *Hs*Prx3-6his through coordination to Ni^{2+}. The switch from high molecular weight stacks to low molecular weight rings was achieved with the addition of chelating agents to the solution.

Throughout this study, novel protein building blocks have been developed that assemble and disassemble in a controllable manner in response to variations in environmental conditions. The assembly system of an existing protein tecton (*Hs*Prx3-WT*)* has been enhanced, creating a protein building block that associates within the physiological pH range. These new routes towards controlled protein oligomerisation could be utilised in future work to form protein nanomaterials for specific functions.

Chapter One: Introduction

Overview

Proteins are emerging as attractive tools in nanotechnology for the formation of functional and responsive materials. The methods developed in supramolecular chemistry have been incorporated with the innate self-assembly ability of biological molecules to create a new field of science known as bionanotechnology. The following chapter will discuss how the field of bionanotechnology has advanced over the last 20 years to develop complex self-assembling structures such as nanotubes [1], layers [2] and cages. [3] The proteins used in this study will also be described in detail to explain how they may be used as novel protein building blocks.

1.1 Nanoscience

1.1.1 What are nanomaterials?

Nanomaterials are generally defined as materials that have at least one dimension between 1 - 100 nm although there is still some debate over this classification. [4, 5] In comparison to bulk materials, nanomaterials have a large fraction of surface atoms per unit mass [6] which means that they exhibit size dependent characteristics and properties that are intrinsically linked to their small dimensions. [5, 7, 8] The stark increase in surface area in comparison to the bulk material means that the inter-subunit forces are far greater than those of gravity and inertia. This leads to the formation of particles that are thermodynamically unstable or metastable and one of the big challenges in the field of nanoscience is developing particles that do not grow in size.

The field is widely interdisciplinary due to the unique properties that these materials offer with areas of research in optics, electronics and materials science.[8, 9]. Because of this there has been a marked increase in interest in the field of nanoscience in recent years. [10]

1.1.2 Top-down and bottom-up approach

There are two main approaches to form nanomaterials, these are the top-down and bottom-up approach. The top-down approach involves the reduction in size of physical bodies until they reach nanoscale dimensions. [6] This can be achieved through a variety of different methods such as nanolithography, where surfaces are etched on the nanoscale using atomic force microscopy (AFM). [11] Due to the nature of the methods used, this particular approach brings with it inherent problems. The lack of accuracy, that can sometimes be an issue when using the top-down approach, can lead to imperfections on the surface of the material. [6] As there is a large fraction of surface atoms per unit, any imperfections can have a profound effect on the properties and characteristics of the final product.

The bottom-up approach describes nanomaterials formed through the assembly of units of a lower order. [12] While the materials formed from this method tend to have fewer surface imperfections, [6] there are still problems relating to the metastable nature of many nanoparticles which can make it difficult to prevent continued aggregation. To be able to control the assembly and form functional and finite nanomaterials in a bottom-up manner it is essential to understand the routes by which these structures are formed.

1.2 Self-Assembly

Self-assembly is the process by which ordered structures are formed from separate components in a disordered system. The process is driven by interactions between the components. These interactions are non-covalent or weak covalent interactions [13] such as van der Waals, hydrogen-bonding, electrostatic, hydrophobic and amphiphilic interactions and metal-ligand coordination. The combined interactions need to be strong enough to overcome the decrease in entropy which occurs when the system becomes more ordered. To be able to design components that can assemble in a controllable manner it is important to understand and predict the interactions that take place. This field of science is classified as supramolecular chemistry.

1.2.1 Supramolecular chemistry

In the past decade a lot of work has been done in the field of supramolecular chemistry to explain the mechanisms by which molecules self-assemble. [13, 14] Supramolecular assemblies are often kinetically reversible, this allows for the components to assemble and disassemble until a thermodynamic minimum is reached. [14, 15] The mechanism of self-organisation can be dictated by specific recognition sites on the subunits and the environment in which the assembly takes place. For example, hydrogen bonding may be disrupted in the presence of water whereas a hydrophobic interaction will be favourable. [16] Steric factors also play a significant role in driving assembly with units binding in the order of least steric hindrance. Metal-ligand coordination can be used to drive the formation of products with specific geometries. [17-21] This is particularly useful when designing 3D supramolecular assemblies such as cages as the geometry can be manipulated by the metal centre and the conditions in which the coordination is taking place. [22] While there is some exceptionally progressive research in designing synthetic molecules that can assemble in a tailorable manner, many are still looking to nature as the peak of complex supramolecular chemistry.

1.2.2 Supramolecular self-assembly in nature

"When searching for a parallel assembly technique for atoms and molecules one is well advised to look at the methods of nature."[23]

In nature there is an abundance of examples of self-assembly, from a bacteria swarm or a school of fish to dynamic self-assembly within the living cell. Even the cell itself is comprised of subunits such as fatty-acids, amino-acids and sugars which associate in a complex manner to form the basic unit of life.

There is a huge array of biological molecules that utilise non-covalent bonding to self-assemble. Phospholipids use amphiphilic interactions to assemble into lipid bilayers which form the cell membrane. These essential biomolecules are comprised of a hydrophilic head and a hydrophobic tail which aggregate into their functional component when exposed to water. [24] Single strands of DNA self-assemble via a complex system of hydrogen bonding interactions and hydrophobic stacking which contribute to the double helix structure, discovered by Watson and Crick in 1953. [25] The three dimensional (3D) structure of these molecules give stability to the DNA which is essential for its continued function. [26]

The self-assembly of proteins is an ideal illustration of how nature forms functional, nanoscale devices within living systems [9] using the bottom-up approach. [27] There are three levels of protein assembly that are all intrinsically linked to the function of the protein. The protein assembly is dictated by the molecular composition of the primary sequence. This is a chain of covalently linked amino-acids which make up the polypeptide chain. Secondary structure describes the first mode of protein self-assembly. It is the initial folding of the chain into small 3D conformations. Tertiary structure depicts the interactions of the secondary structural elements which give the protein its geometric shape. Quaternary structure is the interaction between multiple protein subunits to form what is known as a protein oligomer. All three modes of assembly are essential for the protein structure and function.

An example of this in action is the self-assembly of collagen, the most abundant protein in mammals. It is formed through a hierarchical self-assembly process. [28] The basic unit of collagen forms from three polypeptide chains which self-assemble into a triple helix structure. These then assemble into microfibres which continue to combine into strong and tough collagen fibrils *in vivo.* [29] The electrostatic interactions between the repeating sequence of 3 amino acids, Gly-X-Y (where X and Y can be any amino acid but are often proline (Pro) or hydroxyproline (Hyp) respectively),drive oligomerisation and allow for the formation of tough fibrils. [30] These fibrils are found in many connective tissues such as muscles, tendons and ligaments and therefore their strength is integral to their function. [31]

As in the formation of the cell, proteins can interact with other molecules in complex assemblies which are directly linked to the function of the product. A prime example of proteins forming hybrid complexes is biomineralisation, the formation of biominerals within living tissue. [32] There are many different types of biominerals found across all forms of life, silicates in algae, carbonates in invertebrates and calcium phosphates and carbonates in vertebrates, and they are formed for uses as wide as shells in sea creatures to bones in mammals. Their formation is a hierarchical bottom-up process and displays the interplay between organic and inorganic composites. [32]

Proteins can also self-assembly to form sophisticated molecular-machines such as ribosomes [33] and ATP-synthase. [34] Both are comprised of multiple protein subunits that combine via a series of complex non-covalent interactions to generate nanomachines that are essential to life.

1.2.3 Harnessing the power of nature to design objects with specific structure and properties

Through studying how molecules assemble in nature it is possible to design self-assembling nanomaterials from biologics. Furthermore, it is possible to use biological building blocks that have been specifically functionalised via chemical modifications to create materials for a specific task *in vitro*. [35, 36] The materials can be made from multiple different organic and inorganic tectons, as seen in nature, to make hybrid materials that would be difficult to design purely synthetically. [37, 38] Over all, this relatively new field of materials science is known as bionanotechnology. [39]

1.3 Bionanotechnology

Bionanotechnology is an interdisciplinary field of science that uses biological molecules to generate nanostructures with specific functions and properties. [39] In the natural world, biological molecules are assembled in a bottom-up approach through non-covalent and weak covalent interactions between components which drive the formation of supramolecular structures. [9, 40, 41] Unlike entirely synthetic molecules, biological molecules already contain the necessary elements which allow them to interact in a specific way to generate the desired product. [42] As previously discussed, many of these structures have dimensions on the nanoscale and therefore have unique physicochemical properties that are not present in classical materials and systems.

1.3.1 DNA and RNA

Some of the earliest work in the field of bionanotechnology utilised the specific interactions between DNA base pairs to direct the assembly of nanoparticles of specific size and function. [8, 38, 43, 44] Units of DNA can be combined using a "sticky end", a single strand of DNA at the end of the double stranded helix. [45] These single strands can be used like Velcro when different "sticky ends" have complementary arrangements of nucleotide bases with the assembly being driven via hydrogen bonding between the base pairs. [45] The strands are designed so that the final product is the most thermodynamically stable. However, in many

cases there are different conformations that have very similar energy which can lead to a low product yield, due to self-assembly errors, and a heterogeneous mixture of species. [44, 46] One of the most well researched areas of DNA nanomaterials is [47] the formation of 2D and 3D structures via DNA origami. [44, 48-50] DNA origami is the anti-parallel folding of a long, single stranded DNA (~7 kilobase) scaffold onto a set of specifically sized, parallel double helix oligonucleotides (20- 50 nucleotides) that have been aligned to fit a certain geometric shape. [50] These scaffold DNA arrays [50] can be organised into many different shapes (squares, disks, 5-pointed stars) which can then be arranged into larger assemblies. [49, 50] Fu *et al*. formed nanoscale bioreactors using DNA origami scaffolds.[51] A "sticky end" was added to different edges of a rectangle DNA scaffold to form nanoribbons and long and short nanotubes with controllable diameter (fig. 1.1). [51]

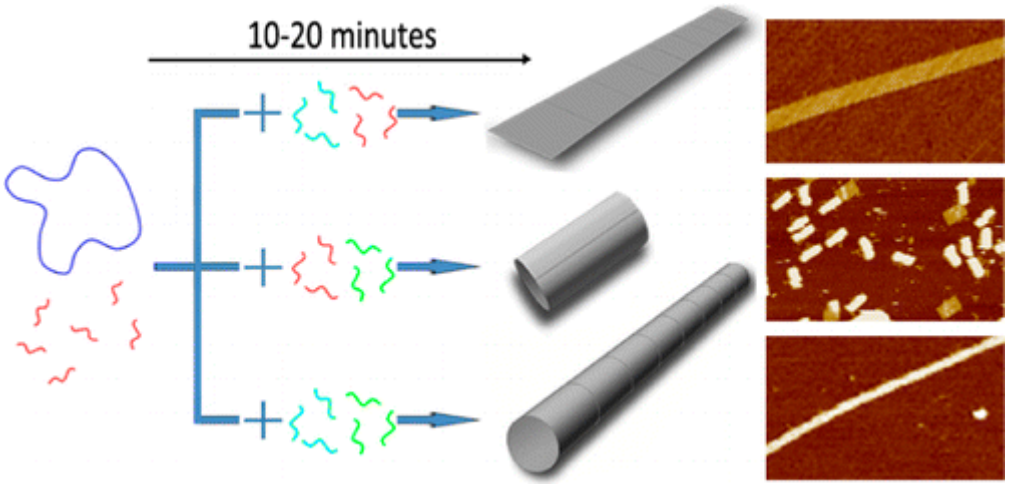


Figure 1.1: Structures formed from DNA origami arrays. The scaffold arrays have been organised into 3D structures using a "sticky end" moiety. [51] *Permission granted*

Another 3D shape formed from DNA is the tetrahedron. A one-pot method for constructing a high yield (~ 95 %) of these shapes was discovered by Goodman *et al.* in 2005. [44] This was a breakthrough in the field of DNA nanotechnology due to the high yield and speed at which the structures formed. The tetrahedron is formed from four rigid DNA helices covalently

joined at one vertex. [44] The rigid strands make up the edges with more flexible branch points at the vertices (fig. 1.2). [44]

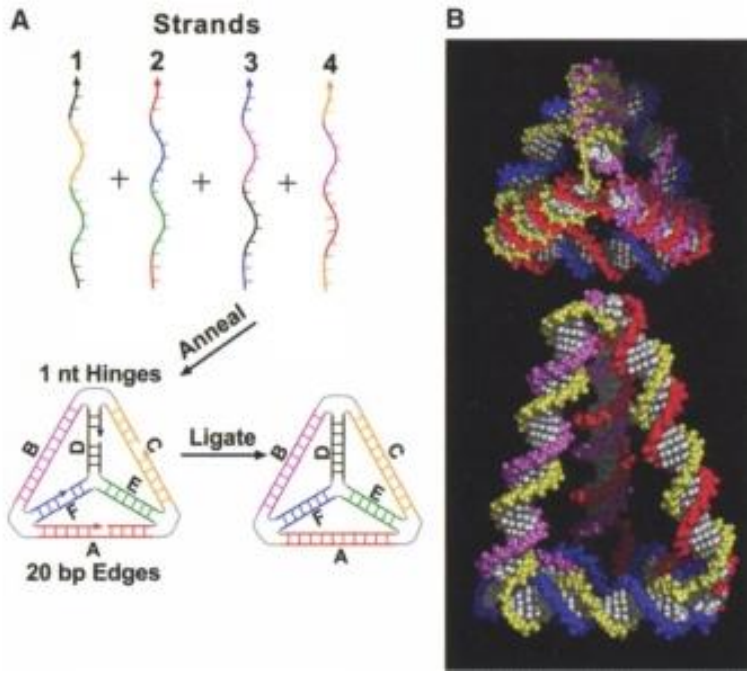


Figure 1.2: DNA tetrahedron formed from four strands of DNA. The rigid DNA helices are covalently linked one vertex with the flexible branch points at the other vertices. [44] *Permission granted*

Ke *et al.* further developed the idea of the DNA tetrahedron by combining the methods of DNA origami and the protocol developed by Fu *et al.* to produced tetrahedral containers with closed faces, something that could not be achieved using the rigid strand and "sticky end" method alone. [43]

The structures formed from DNA can be further functionalised by the addition of components bound to the DNA scaffold. For example, thiolated DNA can be utilised to position gold nanoparticles (AuNP) in an ordered array on the DNA scaffold. [38] It is also possible to anchor enzymes to the DNA scaffold which Fu *et al.* showed can lead to an increase in the enzymatic efficiency when compared to the enzymes freely dispersed in solution.[51]

Ribonucleic acid (RNA) is another biological molecule that can be used to form nanomaterials. [52-55] The nucleotide primary structure can fold via both canonical Watson-

Crick and non-canonical H-bonding into a range of 3D architectures that are believed to be related to the many cellular functions of RNA. [52, 55] By studying the architecture of naturally occurring RNA it is possible to rationally design novel RNA motifs that assemble into a specific shape. [52] For example, the stabilisation of the kink turn (K-turn) motif, found in double stranded RNA (60 ^{o}), [56, 57] by the addition of L7Ae protein to form rigid RNA complexes shaped like equilateral triangles (fig. 1.3). [54] RNA is able to form a variety of 3D structures *in vivo* that are influenced by their environmental conditions (salt content, metal ions) making it an attractive target for the development of biologic-based nano-building blocks. [52] However, materials formed from both RNA and DNA tend to be rigid in nature and lack the dynamics and in-built functionality afforded by other biological molecules. [52, 55]

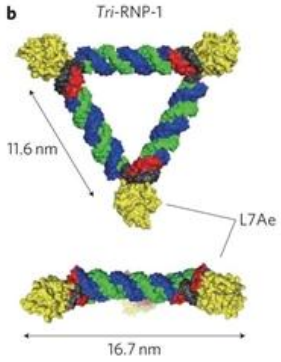


Figure 1.3: RNA equilateral triangle. The K-turn motif is stabilised by L7Ae protein to form a rigid RNA-protein complex. [54] *Permission granted*

1.3.2 Peptides

Peptides offer a range of routes towards self-assembly into ordered structures due to the variety of functional groups available. Specific amino acids can be chosen for the non-covalent interactions they afford. These have been tailored for the design of an array of different structures for example, fibrils, [58, 59] layers, [60, 61] 2D tapes [62] and tubes. [63]

Cyclic peptides have been used since the early 90s to make 1D hollow nanotubes known as cyclic peptide nanotubes (CPN).[64] These nanotubes have been utilised to develop a variety of functional materials from trans-membrane ion channels to drug delivery vehicles and more. [65, 66] While their remarkable stability and the exact driving force behind the

assembly remains unclear it is believed that they interact via hydrogen bonding of the peptide backbone to form anti-parallel and/or parallel β-stacks. The dimensions and surface properties of the tubes can be manipulated by the primary amino acid sequence. In a study by Hourani *et al.* [67] a 3-amino-2-methylbenzoic acid was incorporated into a D, L-α-alternating sequence of a cyclic peptide. This allowed for surface functionalisation from the benzyl group and also reduced the tube diameter due to the methyl, hydrophobic group, being directed into the interior of the ring. The side chains of the amino acids can also be selected to direct modes of assembly. For example, Qu *et al.* [68] used computational methods to study the difference between two cyclic peptides, cyclo[$(_L$-Phe-$_D$-Ala)$_4$] and cyclo[$(_L$-Ala-$_D$-Ala)$_4$]. Steric hindrance and electrostatic repulsion between the phenyl groups on cyclo[$(_L$-Phe-$_D$-Ala)$_4$] in parallel stacking models are reduced when the anti-parallel model is selected. This is not seen for the cyclo[$(_L$-Ala-$_D$-Ala)$_4$]. This demonstrates how the assembly mechanism can be manipulated by changing the amino acid make-up.

Peptides have been designed to form tapes, ribbons, fibrils and fibres via hierarchical systems that are mimetic to proteins in nature. [59, 69] Many peptides are designed to promote β-strand formation which allows for the control assembly into amyloid fibrils. [70] The idea originated from the β-sheet structure seen in misfolded, insoluble proteins in pathological diseases such as Alzheimer's. [71] Aggeli *et al.* developed two 11 amino acid peptides that differed in their side chain content (P_{11} – I –glutamine polar side chains, P_{11}-II phenylalanine and tryptophan aromatic side chains). These small variations in the primary peptide sequence allowed for a different response to peptide concentrations, with P_{11}-II readily forming ribbons, fibrils and brittle hydrogels upon increasing peptide concentrations. [59] The hierarchical assembly of collagen has also been used as a basis for peptide amino acid content to form collagen like fibres. [69, 72] As described above, collagen forms from three amino acid chains which self-assemble into a triple-helix. These chains then assemble to form fibres which grow both linearly and laterally. By using the repeating amino acid sequence of Pro-Hyp-Gly, O'Leary *et al*. was able to develop peptides that would mimic this process in vitro to form hydrogels (fig. 1.4). [72] The nucleation of the fibres was further probed by Sarkar *et al.* and it was found that large nucleation domains undergo rapid fibre formation which leads to gelation or precipitation. [69]

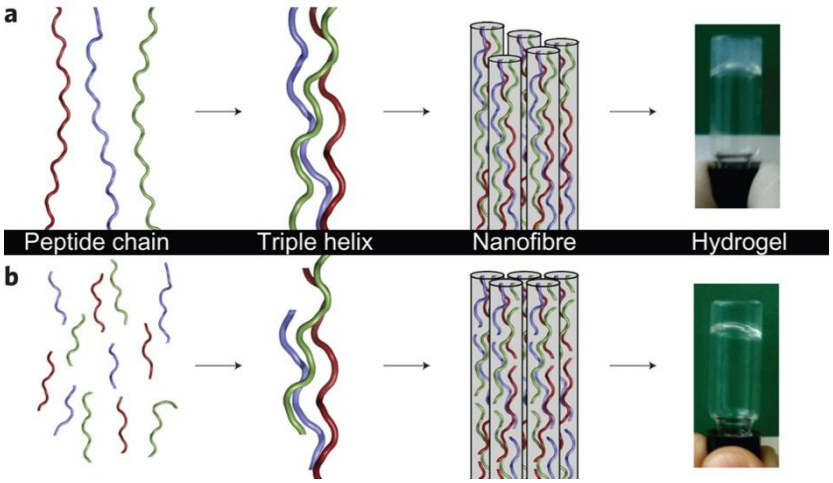


Figure 1.4: Peptides derived from the collagen sequence (a) The repeating amino acid sequence, Pro-Hyp-Gly (b), enables the hierarchical assembly of peptide fibres that form hydrogels. [72] *Permission granted*

A study by Valery *et al.* showed that peptide sequences derived from β-continuous interfaces of homo-oligomeric proteins were ideal candidates for forming peptide nanomaterials.[73] Four peptides, 7 – 8 amino acids in length, were made from the protein interfaces of bovine peroxiredoxin 3 homodimer, bovine β-lactoglobulin homodimers, *Mycobacterium tuberculosis* diaminopimelate decarboxylase homotetramer and *E.coli* umud' protein homodimers.

All four peptides formed birefringent hydrogels from complex nanostructures ranging from bundles of nanofibrils, single nanofibrils and nanoribbons. Circular dichroism and ThT assays confirmed the presence of β-sheet arrays and, interestingly, the self-assembly was reversible, with the degree of oligomerisation being sensitive to the peptide concentration. The nanoribbons made from peptides derived from the dimer interface of bovine peroxiredoxin 3 (*Bt*Prx3) have been further studied and it was found that they can be utilised to encode the assembly of perylene imide based organic semiconductors. [62] This is an elegant example of the types of hybrid materials that can be formed using the natural propensity of biological molecules to self-assemble in a controllable manner.

Structures formed from peptides tend to be more stable than protein structures as they can withstand physicochemical conditions that would denature most proteins. [74] However, they are expensive to synthesise on a large scale and also lack the surface functionality that proteins display.

1.3.3 Proteins

Proteins in biological systems form a diverse array of 3D structures with a wealth of functionalities. [39] They undertake a number of cellular roles for example protein synthesis, [33] chromosomal replication, [75] regulated membrane channels, [76] and catalysts for various different reactions. [77] The architecture, physical properties and mesoscopic size are all related to the function of the protein meaning that there is the potential to form functional nanomaterials by utilising protein building blocks that already have a unique inbuilt function. [78] The challenge now is to develop designer protein nanomaterials of specific size, shape and function by either mimicking nature directly or using computational methods to design novel proteins or protein-protein interactions. [79] As there are no changes to the covalent structure of a polypeptide during folding there is a wide range of possible 3D conformations a disordered polypeptide chain can take. [80] This makes it difficult to specifically design novel proteins that assemble in a predictable manner *in vitro*. A solution to this is to modify naturally occurring proteins to mimic the highly precise hierarchical self-assembly seen in nature to develop novel protein structures *in vitro*.[74] Proteins are particularly attractive tools in nanoscience due to their ability to undergo bottom-up synthesis of biochemical materials with atomic level features inbuilt into the design. [81] To be able to program the assembly of protein building blocks it is imperative to have a good understanding of the interactions that drive this process in order to control it.[81]

1.4 Protein as building blocks for nanotechnology

1.4.1 Protein sequence and folding

Primary sequence

The primary sequence of a protein is made up of covalently bonded, naturally occurring amino acids. The amino acids are joined together via peptide bonds that are formed by a polymerisation reaction that occurs within the ribosome. [33] The amino acid content and the order in which they are placed dictate the spontaneous folding of the protein which nucleates the final protein conformation. [82] Therefore, the protein shape and function is encoded within the primary amino acid sequence. [27] This can be used to predict how changes in the amino acid sequence will influence the final protein structure. [37, 83, 84]

Secondary structure

Protein folding is a fundamental example of biological self-assembly. [71] It is the process by which the polypeptide chain folds in to its native, stable conformation. Almost all native protein folded states are the most thermodynamically stable states under physiological conditions. [71, 80, 85] Local folding of the chain is known as the secondary structure of the protein. The most common protein folds are the α-helix and β-strands, both of which are stabilised by hydrogen bonding between the amide and carbonyl groups on the polypeptide backbone. The α-helix is the most common protein secondary structure (around 40 %). [71] It has 3.6 residues per turn, and the average length is around 10 residues. [86] It is almost always right handed in proteins with all hydrogen-bonds pointing in the same direction. [87] β-Sheets are composed of nearly fully extended polypeptide chains known as β-strands. Interactions via hydrophobic and electrostatic forces are believed to drive the formation of this secondary structural conformation. [27] There are two main types of strands; anti-parallel and parallel (fig. 1.5). [88]

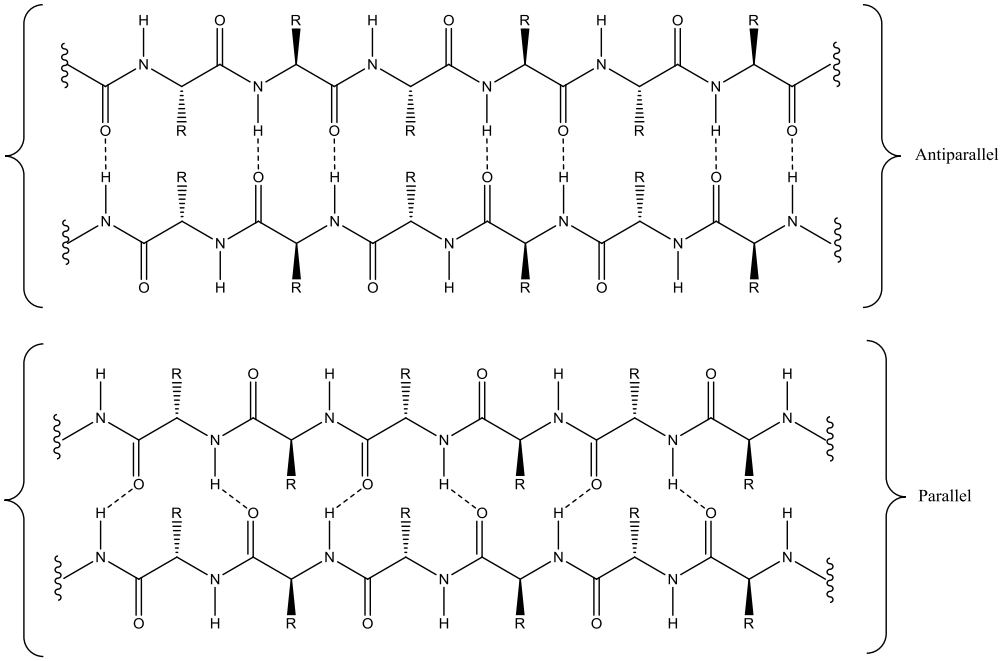


Figure 1.5: The two main types of β-strands.

Secondary protein folding is a crucial factor that affects the overall structure and function of the protein. [88] This can be seen when misfolding leads to the formation of non-native

protein aggregates such as amyloid fibrils. [71, 89] These are seen in disorders such as Alzheimers and Parkinson's disease. [27, 71] Furthermore, by mimicking these secondary structural elements at the protein-protein interface, interactions in protein oligomers can be stabilised. [87]

Tertiary structure

Tertiary structure is the overall shape on the protein, (globular, rod-like, sheet-like). The geometric structure is often related to the function of the protein for example the formation of active sites for enzymes[90] or the 3D structural elements that enhance protein-protein interactions to form biological "nanomachines". [33] It is important to have an idea of the tertiary structure of the protein as the molecular information for assembly can be encoded within the protein's 3D structure. [91]

Quaternary structure

"Proteins in biological systems rarely act in isolation." [90]

The quaternary structure describes the oligomerisation of multiple protein subunits to form the native oligomeric state. The process is driven by molecular recognition, one of the most important chemical events occurring in biological systems. [92] For protein-protein interactions to occur there needs to be complementary surface residues on the separate protein units. Hydrophobic interactions bury non polar residues within the interface (away from aqueous surroundings) which can be further stabilised by Van de Waals interactions. As well as Van de Waals interactions between residues there is an increase in Gibbs free energy due to the expulsion of water from the interface. Electrostatic interactions occur between charged residues on the protein subunit interface that can combine due to complementary charge stabilisation and the formation of salt-bridges. Hydrogen-bonding can occur at the subunit interface via bonding along the backbone of the polypeptide chain. This can be in the form of β-sheets and α-helices, mimicking the secondary structural elements described above. Covalent interactions can also develop in the form of disulfide-bonds. Many protein-protein interfaces have a combination of these interactions which ensures stability.

1.4.2 Manipulating the protein-protein interface

Naturally oligomerising proteins can be the inspiration for the formation of novel assemblies. With an understanding of how these form and function in nature it is possible to derive functional nanomaterials by manipulating and modulating how these supramolecular assemblies aggregate *in vitro*. [81] The features that instigate oligomerisation are already encoded in the protein sequence and the goal is to be able to control the aggregation in a programmable manner by varying environmental conditions and/or site specific mutations.

Interface surface area

Many proteins' interfaces are bound together by relatively weak non-covalent interactions. The stability of multi-subunit complexes is often derived from the large surface area of these interfaces. A recent study completed by Chen *et al.* showed the direct correlation between the binding affinity of 113 heterodimers and the interfacial buried surface area.[93] They concluded that increasing the area of the buried surface leads to an increase in the binding affinity between the proteins. Within the study it was also noted that there is a decrease in the amount of free-energy per unit area with an increase in the area of the interface. (13 cal mol^{-1} $Å^{-2}$ for interface of around 694 $Å^2$ reaching a basal average of 4 cal mol^{-1} $Å^{-2}$ at 2000 $Å^2$). [93] This agrees with the idea of hot-spots within the buried surfaces with the fraction of the surface residues that make energetic contributions decreasing with an increase in total surface area. [93] Alanine scanning mutagenesis is used to analyse hot spots at protein-binding interfaces. Substituting a "hot-spot" residue for an alanine results in a drop in the binding constant of the protein. [94] These sites, when identified, could be targeted for the promotion of oligomerisation.

Site specific mutations

Site specific mutations can be used to further stabilise native interfaces, produce switchable oligomers or even drive the assembly of non-native interfaces. An example of this in action is the addition of cysteine residues to the top and bottom of the ring shaped protein HcpI. This allows the formation of protein nanotubes via a "non-native" ring-interface, stabilised by covalent disulfide bonds. [95] The interfaces can be switched "on and off" with the addition of reducing agents that break the thiol-thiol interactions.

Metal-directed oligomerisation can be utilised by substituting in a residue with a metal coordinating side group, for example, histidine. This coordination is frequently utilised to stabilise protein-protein interfaces. [22, 83, 96-99] The addition of a ligand to a potential protein-protein interface has been shown to drive the assembly of a native monomeric protein, the four helix bundle haem protein cyt *cb*-526, into 1D tubes and 2D sheets via coordination of the histidine side group imidazole, which has been placed to optimise assembly with Zn^{2+}. [96] The resulting assemblies are exceptionally stable and can be used as a template for the formation of Pt^{0} nanocrystals. [97] Additional control over the geometry of the protein oligomers can be gained by analysing the geometry, specificity and affinity of the metal ion-ligand coordination [100, 101] which gives an extra level of protein nanomaterials design.

Histidine residues have also been utilised as pH responsive residues at protein-protein interfaces.[102] The pK_a of the imidazole side group can vary but is generally around the physiological pH range. The subsequent protonation at pH below the pK_a can lead to beneficial [103] or detrimental electrostatic interactions [102] at the interface.

Symmetry factors

Symmetry in quaternary structures arises from repeating protein units known as protomers. There can be no inversion centre or mirror symmetry in the oligomer as this will disrupt the chirality of the protein. [104] Padilla *et al*. showed that structures with complex symmetries can be developed using units with only two distinct symmetry elements. [105] They used this idea to generate self-assembling protein cages, layers, crystals and filaments. By fusing together protein A, a self-assembling dimer, to protein B, a self-assembling trimer, they developed bespoke nanohedral particles (fig. 1.6). The symmetry of the protein units is fixed by linking the domains with amino acids that prefer an α-helical conformation, giving the linker rigidity. [105] This can be seen as a cut and paste route towards protein assembly.

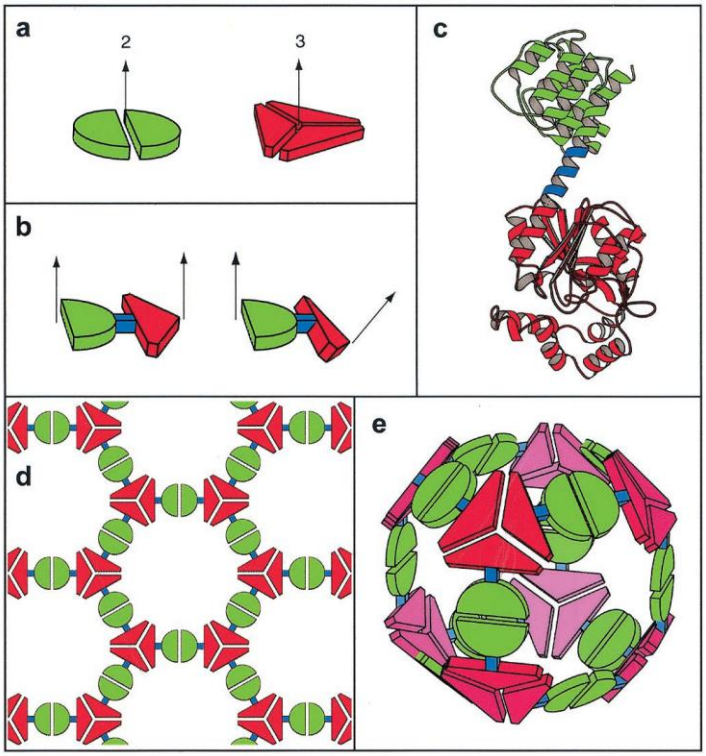


Figure 1.6: Structures with complex symmetries developed from units with only two symmetry elements. Subunits from a dimer and trimer (a) are linked together by a ridgid α-helix to (b and c) to create a tecton that can form sheets (d) and cages (e). [105] *Permission granted.*

1.4.3 Examples of protein nanomaterials formed from naturally occurring proteins

Protein cages

An ideal example of utilising a naturally occurring protein to form a functional nanomaterial is the use of the cage protein, ferritin. Ferritin cages in nature are used to store iron in a microcrystalline non-toxic form. The size of the protein cage controls the size of the non-particulate iron. This knowledge was utilised by Allen *et al.* to design a protein cage that produced cobalt oxide nanomaterials. The ferritin like protein from *Listeria innocua*, a 12 subunit cage with a 5 nm interior, was employed to synthesise size-controlled cobalt minerals. [3] The composition of the mineral could be further manipulated by the temperature in which the reaction took place. It was hypothesised that the acidic residues on the surface of the cage interior were able to mimic biomineralisation nucleation sites found in nature. [3]

Protein Layers

Proteins have also been used to form 2D arrays of layers. One such example was derived from a Class I fungal hydrophobin which self-assembles in nature into amphiphilic monolayers at hydrophobic – hydrophilic interfaces. [2] This protein was use to facilitate the reverse wettability of hydrophobic and hydrophilic surfaces. [2] Coating hydrophobic single walled carbon nanotubes (SWCNT) with a hydrophobin I enables the tubes to be suspended in an aqueous solution. Sonication of SWCNT with $EAS_{\Delta 15}$ (truncated Class I hydrophobin from *Neurospora crassa*) in water resulted in a clear black solution whereas sonication with SWCNT in water alone led to the production of visible particles.

Amyloid fibrils

Protein fibres and amyloid fibrils: make up some of the fundamental building materials of life (structural proteins- collagen, elastin, keratin and spider silk). It is believed that all proteins have the ability to form amyloid fibrils. [71] This can be seen in the lack of similarity between the native structures of amyloid forming proteins. [89] The fibrils tend to have a high β-sheet content and are stable in harsh conditions. [106] Due to the variety in the amino acid sequences of proteins, there is the potential for surface chemical modification of these fibrils, a possible route towards functionalisation. [42, 106]

Amyloid fibrils are believed to be formed from partially unfolded protein intermediates that polymerise into elongated fibrils. [71] This can be achieved under harsh denaturing conditions e.g. acidic pH (pH 2) and heat (80 ^{o}C) and the subsequent nucleation of the denatured protein over time at room temperature. [42] Sasso *et al.* used the dairy industry by-product, whey protein isolates, to form amyloid fibrils in this manner. [106] The fibrils were functionalised via two surface chemical modifications, biotinylation – used for enzyme attachments and thiolation – to anchor the fibrils to a gold surface. Using the biotin-streptavin interaction they were able to successfully anchor both fluorescent quantum dots and glucose oxidase to the amyloid surface. TEM and AFM showed that the amyloid fibrils bound to a gold surface after thiolation using Traut's reagent. This study is an elegant example of how protein nanomaterials can be engineered for a specific function.

Protein rings and nanotubes

Protein rings provide ideal building blocks for the formation of a range of different protein nanomaterials due to the regular symmetry and high surface area of the ring-ring interface.

[107] These can be used as single rings, stacks, tubes, cages and cantenanes. The rings can either be the native oligomeric state or be created by designing non-native interfaces. For example, Bai *et al*. utilised a rigid dimer, *Schistoma japonicum* gluthanione S-transferase (*sj*GST) to form ordered ring structures through metal-ion coordination and non-covalent interactions. [101] By placing an extra histidine, perpendicular to the C_2 axis and adjacent to a native histidine, they were able to use the geometry of Ni^{2+} coordination to imidazole to join the dimers in a V-shape, so that they assembled into a ring. Through computational modelling, they identified key residues at the inter-dimer interface that were mutated to form non-covalent electrostatic interactions, further stabilising the ring form. In fact, the geometry of the ring could be manipulated by disrupting or enhancing these electrostatic interactions by varying the salt content of the solution in which they were formed. [101]

Human peroxiredoxin 3, a naturally occurring ring shaped dodecamer made up of 12 subunit monomers, can be assembled into tubes at acidic pH (pH 4.2) (fig. 1.7). [108, 109]

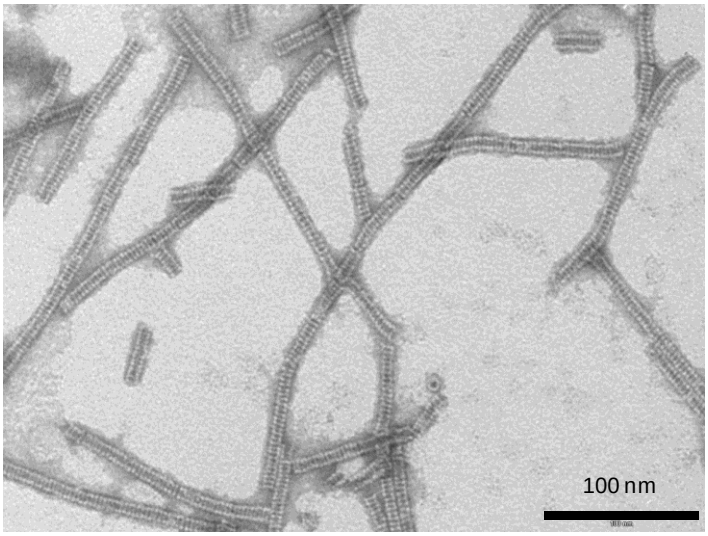


Figure 1.7: TEM image of *Hs*Pr3-WT tubes formed at pH 4.2 Lowering the pH induces the structural changes needed to facilitate interactions at the R-type interface. [110]

Phillips *et al.* utilised cryo-EM imaging to generate 3D models of the tubes into which they docked the crystal structure of a double stacked peroxiredoxin decamer [110] to identify the points of contact that were stabilising the ring-ring interfaces. [109] In accordance with the

information in the literature [110, 111] and the residues that were in close enough proximity to interact, they surmised that the oligomerisation is being driven by electrostatic interactions from Asp^{20} and Lys^{22} on α2 of one subunit in the base ring with $His^{164'}$ on α6 of the subunit directly above. These contacts are facilitated by the local unfolding of the first turn of the α2 helix and the increased disorder of the C-terminus. Lowering the pH induces these structural changes which brings the interacting residues closer together [110] and therefore enables the electrostatic contacts that stabilise the ring-ring interface. [109]

1.4.4 Good building block

The aim of this thesis is to select proteins that have an innate ability to form elaborate quaternary structures through self-associating interfaces. [104] This can be achieved using naturally oligomeric proteins and experimenting with varying conditions to drive the non-native formation of highly ordered nanoscale materials. The three proteins selected have the capacity to do this and therefore work was done to further control their assembly to form the desired protein nanostructures.

1.5 Lsr2 and Nterm-Lsr2

Lsr2 is a nucleoid associated protein (NAP) found in mycobacteria and related actinobacteria. [75, 112-114] It has been shown to have a range of functionalities such as modulating chromatin organization, compaction and global gene expression. [115] It can also protect mycobacteria against reactive oxygen intermediates. [116] Being a highly conserved protein that is found in all sequenced mycobacteria, also makes it a target for the treatment of pathogenic diseases. [117, 118] It belongs to a group of proteins known as histone-like proteins which silence foreign DNA in bacteria. These transcription regulator proteins have the ability to bind selectively to AT rich chromosomal DNA and share similar mechanisms and functions but have little sequence and structure homology. [113, 114, 119, 120] Lsr2 is comprised of an N-terminal, dimerisation (and possible polymerisation) domain [75] and a C-terminal DNA binding domain. [113] It was selected in this study due to the specific way it can oligomerise via the N-terminal region and the ability to bind to DNA, forming large protein-DNA oligomers. [118] Both of these attributes make Lsr2 an attractive candidate for protein nanotechnology, providing a self-assembling tecton with potential in-built function.

1.5.1 Lsr2 physiological roles

Histone-like proteins exist in prokaryotes and perform similar functions to their eukaryotic counterparts, silencing, organising and regulating chromosomal DNA. [112] While there is little sequence homology between these proteins, they share many of the same physical attributes, namely their small size and high positive charge. [121] Lsr2 has been identified as a histone-like protein not only due to its low molecular weight and alkaline pI, but also its ability to bind to DNA, [68] and its broad down regulating activity. [118]

Lsr2 has been identified as a primary silencer of DNA acquired by horizontal gene transfer. [114] Horizontal gene transfer is the act of transferring genetic material between free living bacteria. [122] The process can be both beneficial, adaptation to new environments, and problematic as it can lead to the inappropriate expression of a foreign gene. [119] Xenogenic sequences, DNA that has been acquired from a foreign source, display a higher adenine (A) and thymine (T) content than the host genome. [123] There is an abundance of evidence that Lsr2 binds to AT rich (and conversely, GC deficient) DNA. [68, 112-114, 119, 124] This is one reason why it is believed that this protein plays a crucial role in silencing and regulating xenogenic DNA. [114, 119]

The bacterial chromosome is ordered into a compacted structure known as the nucleoid. The nucleoid is comprised of topologically independent loops that are compacted together by nucleoid associated proteins (NAPs). [121] Histone-like nucleoid structuring protein (H-NS) is known to complete this task in *E. coli* and Lsr2 has been shown to have a similar role in mycobacterium. [124] The means by which this occurs is believed to be via DNA bridging, whereby two DNA binding C-terminal domains associate with two strands of DNA and the bridging occurs through the N-terminal dimerisation domain. [68, 119, 124] In fact, it has been shown that both of these domains are essential for Lsr2 function *in vivo*. [124] Furthermore, AFM studies *in vitro* display the ability of Lsr2 to compact relatively non-specific regions of DNA. [124]

Lsr2 has been identified as a global transcription regulator due to its ability to regulate gene transcription [125, 126] and its apparent non-specific binding to DNA sequences. [113] This capacity to regulate gene transcription has been shown to be integral to the antibiotic response of *Mycobacterium tuberculosis* (*M. tuberculosis*). A strain containing an increase in Lsr2 expression was more resistant to the antibiotic, ethanbutol (EMB). [118] Furthermore,

Lsr2 knockout cell lines (Δ*lsr2*) showed up-regulation of antibiotic induced genes indicating that Lsr2 acts as a silencer to down-regulate these genes. [118]

During transfer into a host, pathogenic bacteria are subjected to stress conditions. This can range from the change in temperature to the attack of the parasitic cell by the host immune responses. One such immune response is the attack by reactive oxygenated species (ROS). In *Mycobacterium smegmatis* (*M. smegmatis*) Lsr2 has been shown to protect DNA from ROS, with deletion of the *lsr2* gene leading to more susceptibility to the detrimental effects of H_2O_2 DNA damage. The role of protection was demonstrated *in vitro* by the direct binding of Lsr2 to the pUC19 plasmid DNA under stress conditions. [116] For a pathogen to survive transfer into a host, there needs to be a degree of adaptation to the new environment. One significant change is the level of oxygen as the pathogen passes from the atmosphere into the host. The ability of *M. tuberculosis* to adapt to changing oxygen availability is one key reason for its success. [125] Lsr2 deletion leads to a decrease in this ability indicating that it is needed for adaptation to the new oxygen environment. [125] However, unlike *M. smegmatis*, there is no difference to a knockout mutant that has been subjected to ROS stress under normal oxygen conditions. This indicates that the protein is more likely to have a transcription regulation role rather a direct DNA protection role in *M. tuberculosis*. Due to the virulence of *M. tuberculosis* and its tolerance to antibiotics, proteins that are involved in host adaptation are key targets in the defence against pathogenic diseases.

1.5.2 Lsr2 structure

C-terminal DNA binding domain

The structure of the C-terminal DNA binding region was solved by Gordon *et al*. using nuclear magnetic resonance (NMR). [113] It was found that the C-terminal DNA binding domain binds to the minor grove of the DNA double helix. [113, 119] The C-terminal domain is comprised of residues 77 – 137 and exists as two α-helices joined together by a loop. It is unique to other histone-like proteins in the fact that the helices are ordered perpendicular to one another and stabilised in this conformation via hydrophobic interactions (fig. 1.8). [113]

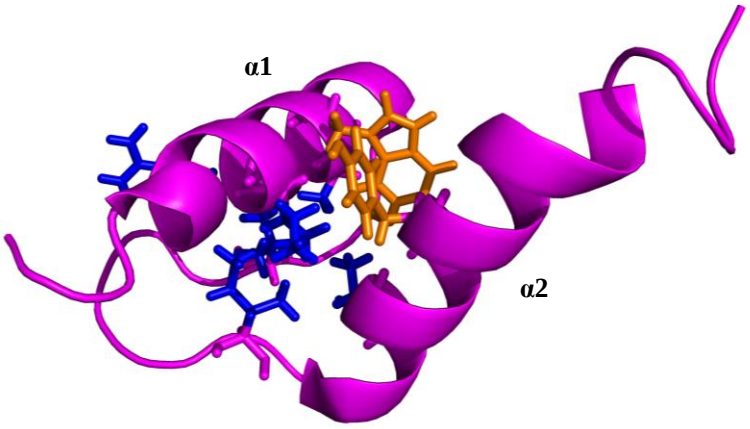


Figure 1.8: C-terminal domain of Lsr2. Two α-helices are arranged perpendicular with respect to each other, stabilised in this arrangement by the hydrophobic core (blue). Aromatic between Trp86 and Tyr108 (orange) adds additional stabilisation.

The key residues involved in DNA binding were identified by chemical shift changes when bound to DNA; Gly^{73}, Ala^{74}, Ser^{80} – Glu^{85} and Ser^{95} interact with the sugar phosphate backbone to increase the binding affinity. The amine hydrogen chemical shift for residues Thr^{96} – Arg^{99} completely disappears in the presence of DNA indicating an intermediate exchange with DNA on the NMR time scale. Furthermore, the DNA binding is diminished with the substitution of Arg^{97} and Arg^{99} for alanine [114] showing that these residues are integral to the DNA binding. The mechanism of binding was described as an AT-hook like motif, similar to that seen for H-NS even though the roles of this two proteins do diverge [68, 114] and there is little structure and sequence homology between them. [114]

N-terminal dimerisation domain

The N-terminal domain of Lsr2 has been shown to be integral to function.[124] *In vitro* testing of increasing amounts of Lsr2 to DNA shows that Lsr2 can not only bridge DNA through the dimerisation region but can also form bundles of DNA, presumably through the N-terminal oligomerisation region. [124] [68] It was this particular function that identified Lsr2 as a H-NS like protein, with the potential to structure the bacterial chromosomal DNA into a nucleoid. [68]

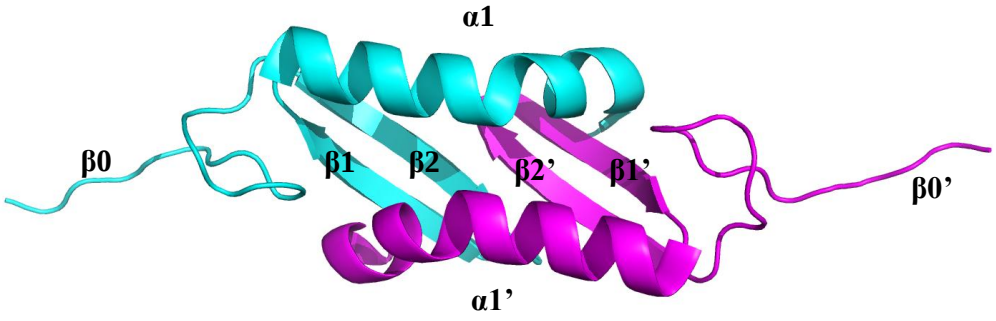


Figure 1.9: N-terminal domain of Lsr2. Dimerisation occurs between β2 on each subunit. Additional oligomerisation at β0 allows the dimers to associate into long protein chains.

In a study by Summers *et al.* the crystal structure of Nterm-Lsr2 (the truncated N-terminal domain of Lsr2) showed a simple dimeric structure comprised of two α-helices, surrounding a four stranded β-sheet, with the dimerisation occurring between two anti-parallel β-strands (fig, 1.9). [75] During the study, it was noted that Nterm-Lsr2 was able to form large oligomers via the disordered N-terminal residues. Closer inspection of the crystal structure showed that three residues, Met^1, Ala^2 and Lys^3 were absent, presumably due cleavage by contaminating proteases. [75] The remaining residues in the disordered N-terminus (KVTVTLV) are well suited for the formation of an inter-dimer β-strand. [73] The oligomerisation can be initiated in a more controlled manner using trace amounts of trypsin (1:500, trypsin:Nterm-Lsr2). The oligomers formed, using this method, are remarkably stable, being visible on both SEC and TEM. [75, 127] Full length Lsr2 and DNA were analysed on DNA agarose electrophoresis in the presence and absence of trypsin. The addition of trypsin leads to the compaction of DNA suggesting the oligomerisation of a DNA-protein complex where the C-terminal region binds to DNA and the N-terminal region oligomerises into larger assemblies, compacting the DNA. [75] This could provide evidence around the mechanism by which Lsr2 acts as an NAP. [75] Indeed, the residues within the oligomerisation region are remarkably well suited to form inter-subunit β-interfaces [73] further suggesting a biological role.

1.5.3 Lsr2 as a tecton in bionanotechnology

It was noted in a study by Colangi *et al*. during [112] electrophoretic mobility shift assays, that Lsr2 appeared significantly larger than would be expected (12 kDa). [118] This indicated the presence of DNA-protein complexes, bound by Lsr2 in a non-specific manner. [118] A further study by Qu *et al*. elucidated a possible mechanism by which Lsr2 is able to perform its physiological roles. [68] Not only can Lsr2 bridge DNA [124] but it can also stiffen both unfolded and folded DNA into a nucleoprotein filament. [68] Steric confinement due to high protein concentration was ruled out by studying the integration host factor (IHF) protein, which binds to DNA as a dimer but did not have the same stiffening effect as Lsr2 at equal concentrations. [68] This knowledge combined with the fact that the N-terminal domain of Lsr2 has the capacity to oligomerise into large structures (fig. 1.10) [75] and also its behaviour as a NAP, organising and shaping DNA, [113] means this protein has the potential to be used in a controlled manner to form specifically designed nucleoprotein complexes *in vitro*. This would join the bionanotechnology fields of bespoke DNA organisation and protein oligomerisation with Lsr2 being used to build functional, nanoscale biological materials. Furthermore, a greater understanding of the behaviour of Lsr2 could be beneficial to the design of protein targeted pharmaceuticals for the treatment of virulent pathogenic bacterial infections such as *M. tuberculosis.*

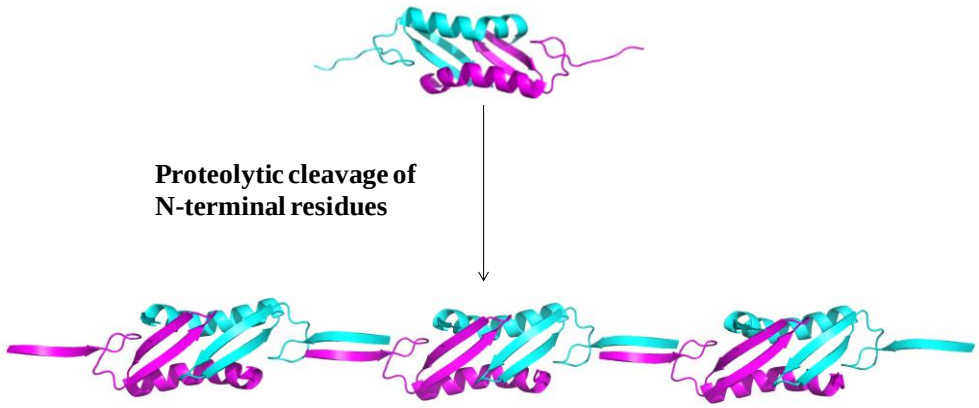


Figure 1.10: Assembly of Nterm-Lsr2 upon the proteolytic cleavage of three N-terminal residues. Removal of these residues allows the formation of an inter-dimer β-strand.

1.6 Peroxiredoxin

Peroxiredoxins (Prx) are a class of peroxidases that exist in all forms of life. [128] Their primary role is to protect the cell against oxidative stress by reducing hydroperoxides, organic hydroperoxides and peroxynitrite within the cell. [129] Unlike other peroxidases, Prx does not contain a redox active metal centre but instead uses a cysteine residue at the active site, conserved in all species. [128] In addition to their peroxidase activity, Prxs are believed to adopt a number of other roles under physiological conditions. For example, as molecular chaperones or intercellular regulators of levels of the H_2O_2 molecules that are involved in redox signalling. [129, 130] There are a variety of different oligomeric states that Prxs can adopt, from oblique dimers, rings of decamers and dodecamers, stacks and tubes and cages. The aggregation of the subunits is dependent on the classification of the Prx and there is strong evidence that the oligomeric state is linked to the protein function. [131] Due to the variety of highly ordered, supramolecular complexes that Prxs can form, this class of enzymes has been identified as an ideal candidate for protein nanotechnology. [109] Because of the apparent innate ability of these enzymes to form high molecular weight (HMW) structures, it is proposed that there are further routes towards assembly that are yet to be explored. These routes have the potential to be manipulated to develop tailorable biological nanomaterials *in vitro*.

1.6.1 Catalytic cycle

Unlike other peroxidases, Prx does not contain a redox active metal centre; instead, it completes the catalysis via a redox active cysteine. The catalytic cycle is comprised of three steps (fig.1.11):

(1) Peroxidation
(2) Resolution
(3) Recycling

The active site is organised in a way that allows the binding of peroxides in a conformation for an S_N2 substitution reaction. [132] The environment of the active site lowers the pK_a of the peroxidatic cysteine (Cys_p) side group which gives a fully negatively charged thiolate. This attacks the peroxide to form a sulfenic acid. A resolving thiol, either from the Prx itself or another small molecule, attacks S_pOH, releasing water and forming a disulfide bond. For this to occur, there needs to be a degree of local unfolding at the active site which brings S_P-

OH out of the protected, fully folded active site pocket. The formation of the disulfide bond effectively locks the locally unfolded (LU) conformation of the reaction intermediate. The disulfide bond is reduced by a thiol group from another protein or small molecule. [133, 134] Many Prxs incorporate thiorodoxin or a thiorodoxin like protein for this role. Once the disulfide bond is broken, the fully folded (FF) active site conformation is renewed allowing the protein to be ready to bind to another peroxide substrate. It is possible for S_pOH to react with another peroxide before the resolution step occurs. This leads to formation of a sulfinic acid which cannot react with the resolving cysteine, leaving an inactive form. However, the overoxidised cysteine can be recycled back to its active form via an ATP-dependent reaction with sulfiredoxin (Srx). The overoxidation of the peroxidatic cysteine is believed to be one of the ways that Prx can be involved in additional physiological processes. [135]

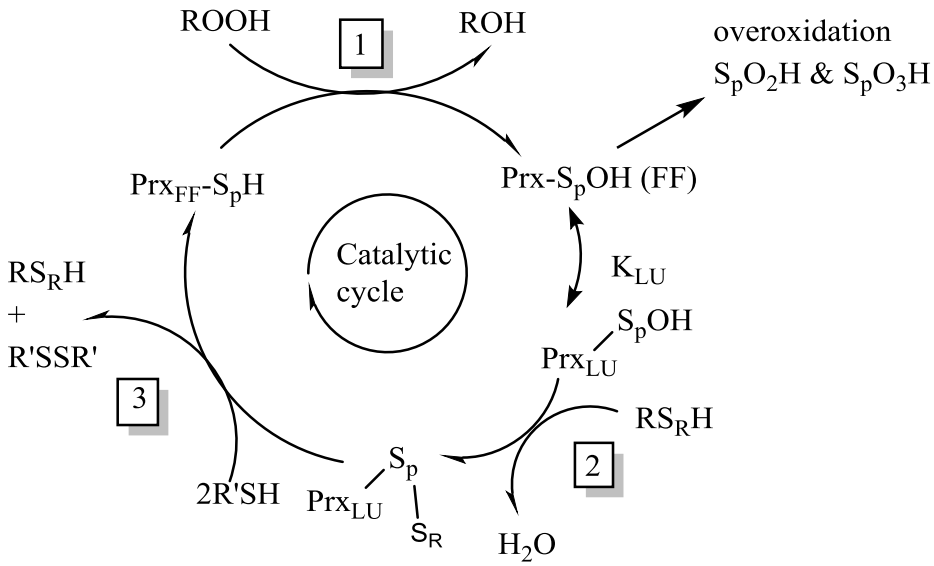


Figure 1.11: Peroxiredoxin catalytic cycle. 1: Cys_p reacts with the peroxide substrate to form sulfenic acid. 2: Local unfolding of the active site proceeds attack by Cys_r to form a disulfide bond which releases water. In some cases another peroxide can react with the sulfenic acid to form a sulfinic acid. Disulfide bond formation locks the LU unfolded conformation. 3. Disulfide bond is reduced by thioredoxin or another thiolate-bearing reductant which renews the FF active site conformation.

1.6.2 Peroxiredoxin classes

Until relatively recently Prxs were grouped into three main classes of, typical 2-Cys, atypical 2-Cys and 1-Cys. 2-Cys Prxs have two cysteine residues that are involved in the catalytic cycle, the peroxidatic cysteine - Cys_p and a resolving cysteine -$Cys_{r,}$ which regenerates the reduced form of Cys_p. [77] In the typical form, Cys_r is located on the adjacent subunit of the oblique dimer. Atypical 2-Cys Prxs, however, form intra rather than intermolecular disulfide bridges in the reaction intermediate. [77] 1-Cys Prxs have only one enzymatic cysteine and require another small molecular or protein to donate the resolving cysteine. [77]

The issue with this classification is that it only describes one aspect of the Prx roles and many Prxs that have the same 1-Cys, 2-Cys, typical or atypical classification vary hugely in both sequence and physiological behaviour. [136] For this reason, a bioinformatics approach called active site profiling was adopted to analyse the sequences of segments situated near to the active site of 29 well characterised Prxs. [137] This information was then extrapolated to identify ~3500 known Prx sequences and the enzymes were classified as described in the table on page 28 (Table 1.1). [136] The protein used in this study is a typical 2-Cys Prx from the Prx1 subfamily. It is found in the human mitochondrial matrix and has been labelled as human Prx 3 (*Hs*Prx3). [138]

Subfamily	Phylogenetic distribution	Structural distinctions relative to Prx core fold	Oligomeric states and interfaces	Typical location and conservation of Cys_r (when present)
Prx1	Archaea, bacteria, plants, and other eukaryotes	Extended C terminus	B-type dimers, (α2)5 decamers (and rare (α2)6 dodecamers) through A interface	C terminus of partner subunit (~99%)
Prx6	Archaea, bacteria, plants, and other eukaryotes	Long, extended C terminus	B-type dimers, some (α2)5 decamers through A interface	No Cys_r (≥41%)
AhpE	Bacteria	Extended loop at N terminus	A-type dimers	Helix α2 (~67%)No Cys_r (≥19%)
PrxQ	Archaea, bacteria, plants, and fungi (not animals)	Extended helix α5	Monomers and A-type dimers	Helix α2 (~61%)Helix α3 (~6%)
Tpx	Bacteria	N-terminal β hairpin	A-type dimers	Helix α3 (>95%)
Prx5	Bacteria, plants, and other eukaryotes (not archaea)	Pi helix insertion in α2i; ~20% fused with Grx domain	A-type dimers	Helix α5 (~21%) Between β1 and β2 of N terminus (~17%)

Table 1: The 6-subfamilies of Prx. Table from Poole *et al*. describe the different Prx subfamilies and how they are classified. [136] Many contain a mixture of typical and atypical 2-Cys and 1-Cys Prxs which is why a new classification method was needed to prevent the over-annotation of newly discovered Prxs. [136, 137]

1.6.3 Peroxiredoxin monomer

The Prx monomer is globular protein that has a common core tertiary structure that is based on the thioredoxin (Tpx) fold. [77, 131, 139, 140] The Tpx fold is based on a central twisted β-sheet which contains a 4 stranded β-sheet (2 strands parallel, 2 stands anti-parallel) surrounded by 2 α-helices (fig. 1.12 A). In addition to N-terminal and C-terminal extensions,

Prxs have a central twisted 5 stranded β-sheet comprised of β5- β4- β3- β6- β7. The sheet is sandwiched between β1- β2-α1 on one face and α2, α3 and α5 on the other (fig. 1.12 B). In the fully folded conformation, Cys_p is in the first turn of the α2 helix. For successful function, the active site needs to be stable in both FF and LU conformations. In a comprehensive review by Hall *et al*. this stabilisation was described as a "cradle" from other secondary structural elements with β3 and β4 acting as the base and α3 and α5 acting as the walls of the cradle. [131] This tertiary fold is highly conserved across the Prx family and therefore has great significance.

A: Thioredoxin monomer **B: Peroxiredoxin monomer**

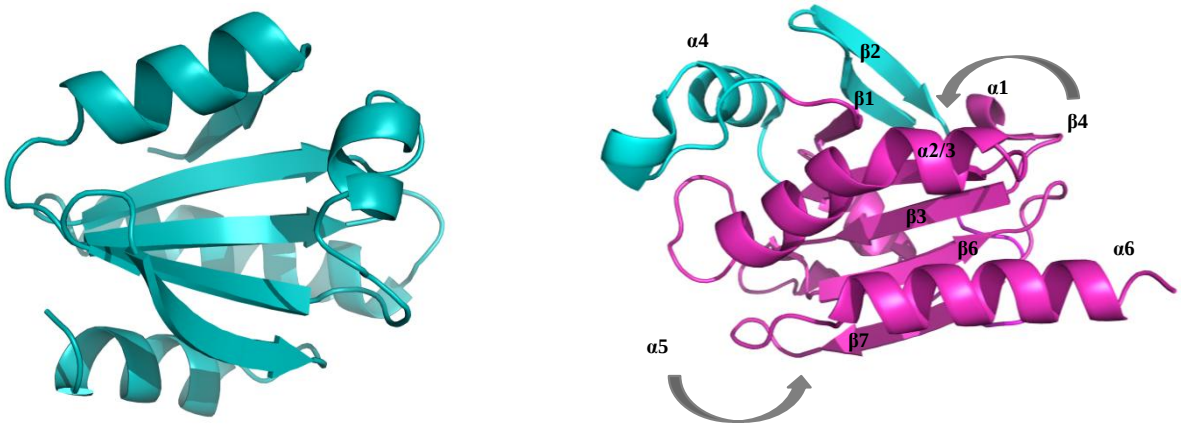


Figure 1.12: Tpx fold and Prx monomer. (A) Thioredoxin is comprised of a central twisted β-sheet surrounded by two α-helices. The Tpx fold bakes up the Prx common core tertiary structure. *Image generated using Pymol with coordinates from PDB:1X0A* [140] (B) The peroxiredoxin monomer with the key secondary structural elements labelled. *Images generated using Pymol with coordinates from PDB:4MH2* [143]

1.6.4 Active site

The active site contains three amino acids which are absolutely conserved[136], these being the peroxidatic cysteine, a proline and an arginine residue. In addition there is also the highly conserved threonine, which is replaced with a serine in some Prxs.[77, 131, 136] The $PXXXTXXC_p$ residues make up the C_p loop with arginine being positioned so that it can reduce the pK_a of the cysteine thiolate via hydrogen bonding stabilisation (fig. 1.13). [77,

138] The hydrophobic proline protects the active site pocket from water and the fully folded active site allows for the positioning of the peroxide substrate to undergo a S_N2 reaction during the peroxidatic step. [132, 141] In typical 2-Cys Prxs, the active site undergoes a local unfolding event that allows Cys_p and Cys_r to adjust position so that they are close enough to form a disulfide bond.

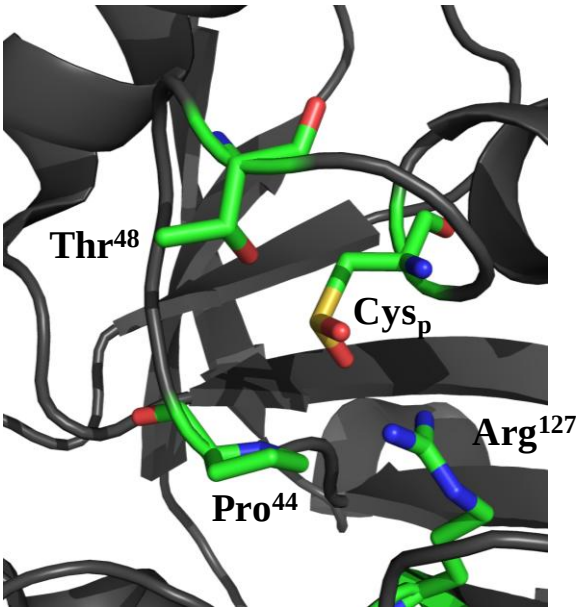


Figure 1.13: Prx active site. The conserve residues are in green. The Cys_p thiolate is stabilised via H-bonding with Arg^{127} and Pro^{44} shields the active site from water. The fully folded active site positions the peroxide substrate for a S_N2 reaction. *Images generated using Pymol with coordinates from PDB:1QMV [142]*

This involves the unwinding of the α2 helix and the disordering of the C-terminus. [143] Additional variations at the opening of the active site have been identified as selective for different forms of peroxide substrate. For example, some Prxs contain hydrophobic residues from the dimer partner that create a hydrophobic collar at the active site. This helps to increase binding efficiency and therefore improve the Prxs rate of reaction with organic peroxides. [129, 136, 144]

1.6.5 Peroxiredoxin dimer

Although there are known Prx monomers, the only active forms of which are found in the BCP subfamily, [134] most Prxs form more complex quaternary structures. [131] The smallest of these is the Prx dimer. Prxs have two potential dimerisation interfaces, the B-type interface, so called due to its β-sheet structure and the A-type interface, alternative or ancient interface.

B-type interface

The B-type interface is formed by anti-parallel interactions along β7 on two monomeric subunits to form a 10 stranded β-sheet. [134] The interaction is stabilised by hydrogen bonding and the burying of hydrophobic residues along the interface (fig. 1.14). In typical 2-Cys Prxs, Cys_r is located across the B-type interface at the C-terminal end of the subunit partner. [145] Upon the unfolding of the first turn of α2, a disulfide bond forms at this interface producing an exceptionally stable dimer that buries around 30 residues per monomer. [146] Prxs that adopt this dimerisation do not disassociate under physiological conditions. [128]

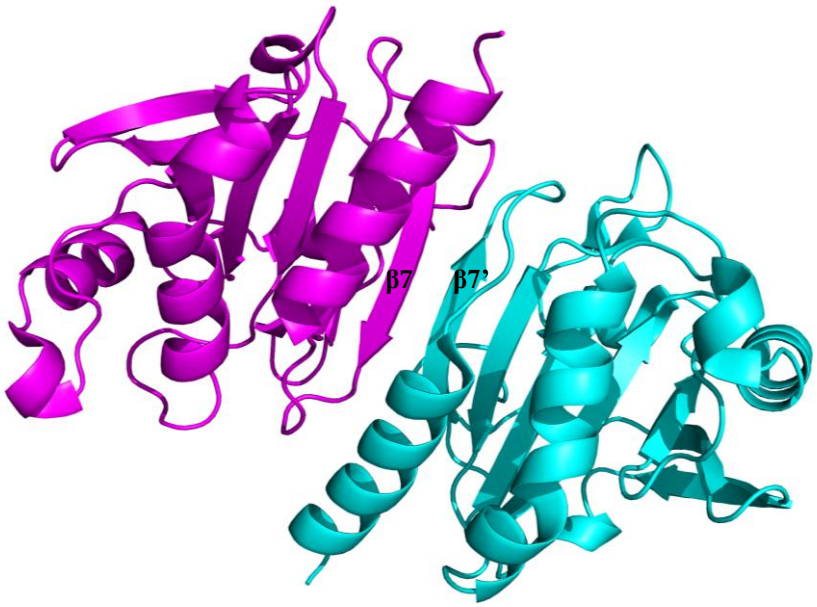


Figure 1.14: B-type interface. The dimer associates along β7 of each subunit to form a 10 stranded β-sheet. This is stabilised by both H-bonding and hydrophobic interactions. *Images generated using Pymol with coordinates from PDB:4MH2*

A-type interface

The A-type interface is comprised of a tip-tip association of chains between β1 and β2' and the loops proceeding α2, α3 and α4. The interface incorporates a mix of hydrogen bonding and hydrophobic interactions for stabilisation. The crystal structure of *Plasmodium falciparum* antioxidant protein (*Pf*AOP), displays a dimer that associates along the A-type interface.[147] The researchers noted 5 regions of interaction between subunits (numbering from *Pf*AOP);

0- residues 11 – 14 and 34
1- residues 54 – 57
2- residues 8-100
3- residues 110 -113
4- residues 127 – 134

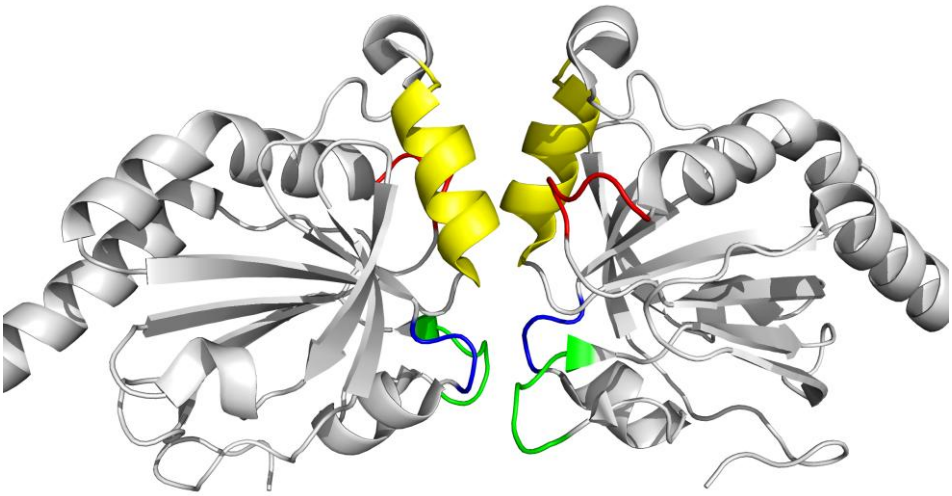


Figure 1.15: A-type interface. [146]. Region 1 (red) precedes the C_p loop and interacts with region 2 (yellow) through hydrophobic contacts. Regions 3 (blue) and region 4 (green) make up the other cluster of contacts at the A-type interface with Phe^{130} and Asn^{111} believed to the key involved residues. The crystal structure of *Bt*Prx3 is used in this figure as it shares 93 % sequence homology with *Hs*Prx3 and the same four regions of contact at the A-type interface as *Pf*AOP. *Images generated using Pymol with coordinates from PDB:4MH2*

There are three main clusters. Clusters I and III are located on the periphery of the interface with region 0 and 3 on one subunit interacting with region 4 of the adjacent unit (fig. 1.15). The key residues involved are believed to be Phe^{130} and $Asn^{111'}$. The other cluster, II, makes up the core of the A-type interface. The cluster is located around the 2-fold axis and therefore the same residues from each subunit, from region 1 and 2, interact with each other. The interface is comprised of both hydrophobic interactions and hydrogen bonding between the residues within these regions. [147] [148]

When compared to other Prxs that associate in this way, Sarma *et al.* found that the interacting regions were conserved across many subfamilies, with the exception being the interactions of region 0, which is an N-terminal extension specific to *Pf*AOP. [147] Due to the conservation of this site across many Prxs and the increased frequency in comparison to the B-type interface, it is believed that the A-type dimer interface was the ancestral interface for Prx quaternary structure. [131, 147]

1.6.6 Peroxiredoxin toroids

Many Prxs utilise both the A-type and B-type interfaces to oligomerise into toroids comprised of 8, [149] 10, [141] or 12-Prx monomers (fig.1.16). [109] The dimer-toroid inter-conversion is redox, concentration and pH sensitive with the toroid being stabilised when the active site is reduced, the protein concentration is high and the pH is relatively low. Interestingly, the overoxidised sulfinic acid form of Cys_p is also a stable, if slightly less compact, toroid that can only be disassembled upon reduction with Srx. [150]

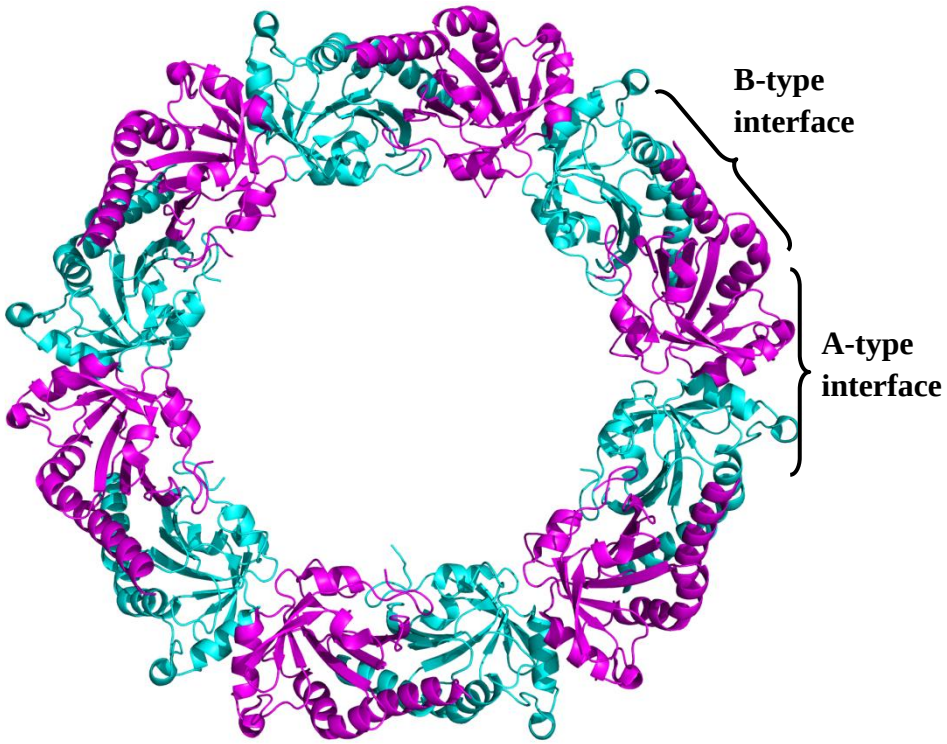


Figure 1.16: *Bt*Prx3 dodecamer. [142] The dimers associate via the A-type interface to form a toroid. *Images generated using Pymol with coordinates from PDB:4MH2*

The Prx used in this study is human Prx 3 (*Hs*Prx3), a mitochondrial Prx found in the matrix, exists as a dodecamer. [109]This is relatively unusual among the known typical 2-Cys Prxs. The bovine analogue (*Bt*Prx3), which has a 93% sequence similarity, is also a dodecamer. [143]

Redox

In many typical 2-Cys Prxs the quaternary structure is influenced by the redox state of Cys_p. [148, 151, 152] The reduced form exists as a stable ring whereas the disulfide bonded moiety can form a range of different oligomeric states depending on protein concentration and pH. By studying the crystal structure of disulfide bonded AhpC and comparing it to other known structures of Prx rings and dimers, Wood *et al.* were able to pin-point key regions that associated at the A-type interface and identify the potential molecular "switch" that dictated the quaternary structure. [148] Unlike other non-reduced 2-Cys Prxs, which crystallised as oblique dimers, the crystal structure of AhpC showed a meta-stable toroid of 10 subunits. It is believed this was due to the high concentration needed for crystallisation which was above the critical range for toroid formation.

Comparisons were drawn between regions 1-4 (above) on the structures of non-reduced AhpC and those of the oxidised dimer – HBP23 and the reduced ring structure of Tpx-B with AhpC being cited as the intermediate of the two forms. It was found that in the reduced state, the C_p loop is less mobile as it is packed against the fully folded α2 helix. Upon oxidation, the α2 helix begins to unfold so that in AhpC the C_p loop is more mobile and therefore there are less interactions in region 1. In the locally unfolded structure of HBP23, the active site has collapsed as Cys_p has vacated the pocket to bind to Cys_r. This leads to the uncoupling of the C_p loop to region 1 which completely destabilises the A-type interface. A later study by Matsumara *et al.* showed that rearrangement of the active site helix led to changes in the positions of key hydrophobic residues at the A-type interface. [152] It was postulated that this could lead the steric clashes at the dimer-dimer interface, further destabilising the toroidal form upon intermolecular disulfide bond formation. The fact that the protein is sensitive to the redox activity that takes part in the catalytic cycle suggests that the oligomeric state of the protein is integral to its physiological function. [77, 128, 133, 145] It has been shown that the final recycling step is more successful when the protein is in the dimeric form. [131] However, the specific reasons behind the toroid form are still unknown as Prx dimers are fully active.

pH

It has long been noted that lowering pH can improve the stability of the Prx 2-Cys ring. [145, 151, 153] Even small variations between pH 7 – 8 have been shown to influence the quaternary structure. [153] Until recently, the specific mechanism that governed this switch

was unknown. Morais *et al*. set out to answer this question by studying the crystal structure of the disulfide bonded *Leishmania braziliensis* Prx 1 (*Lb*Prx1m), at both basic and acidic pH. It was noted that, at pH 4.4, the C_p-loop is locked into a closed conformation due to electrostatic interactions between a protonated His^{113} and Asp^{76}. In addition, there is a hydrogen bonding network that further stabilises this conformation, the positioning of which allows key hydrophobic residues to interact at the inter-dimer A-type interface, helping to sustain the decameric form (fig.1.17 A). [103] At pH 8.5 the protein is in the dimeric state. His^{113} is no longer protonated, removing the key electrostatic interaction with Asp^{76} (fig.1.17 B) and allowing an open C_p-loop conformation, seen in many other disulfide bonded typical 2-Cys Prx dimers. [109, 131, 133, 148]

A: pH 4.0 **B: pH 8.5**

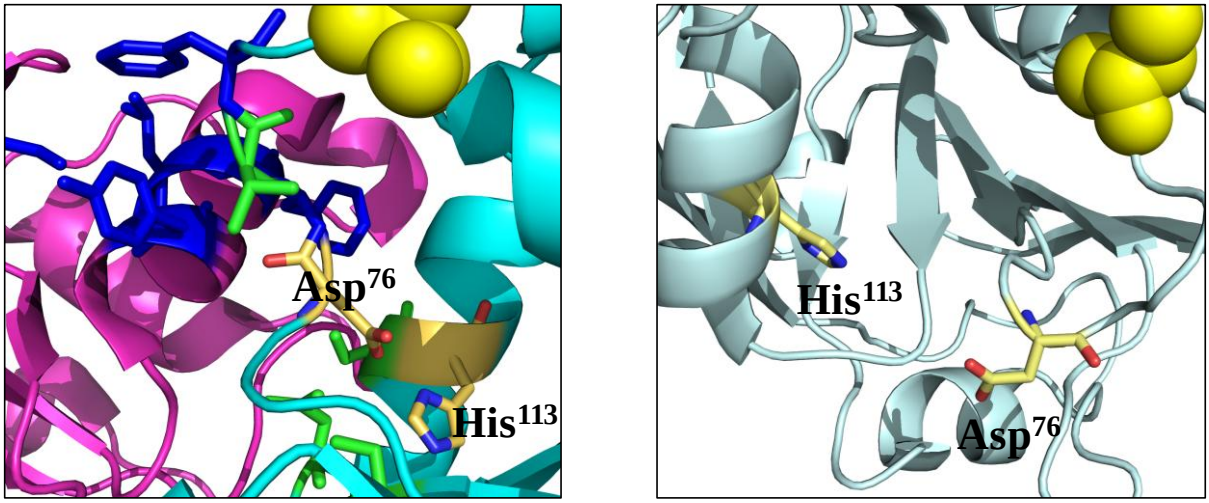


Figure 1.17: pH padlock [103] (A).At pH 4.0 electrostatic interactions between Asp^{76} and His^{113} and a H-bonding network (green residues) lock C_p in a closed conformation. This facilitates hydrophobic interactions (blue residues) at the A-type interface. (B) At pH 8.5 His^{113} is deprotonated. There is no longer an electrostatic interaction with Asp^{76} and therefore the C_p loop is open. This conformation displaces the hydrophobic residues at the A-type interface and the decamer dissociates into dimers. *Images generated using Pymol with coordinates from PDB:4KB3 (A) and 4KCE.*

The solution studies showed that the changes in quaternary structure occurred between pH 7-8, with the pK_a of His^{113} being around 7.37. [103] This is consistent with the hypothesis that the protonation of the histidine was critical for the stabilisation of the decamer in non-reducing conditions. The key residues identified in this study are conserved in many other Prx1 typical 2-Cys Prxs, suggesting that a similar mechanism could be the origin of toroid stability for these proteins as well. [103] Furthermore, the small range of pH that is necessary to facilitate this change suggests that it could play a role *in vivo*, aiding Prx to switch between different physiological roles.

N-terminal tags and extensions

It has been shown in a number of structural studies that additional N-terminal residues can have a profound effect on the quaternary structure of Prx. [141, 151, 154-156]. In a study by Cao *et al.* on the recombinant, histidine tagged Bovine Prx 3 (*Bt*Prx3), it was found that the presence of a tag is enough to keep the ring intact even in non-reducing conditions.[151] A similar effect is seen in solution studies for Myc tagged Tsa1. [154] Wild-type, human PrxIV has a native N-terminal sequence, around 30 residues long, and this human Prx isologue is also a redox stable decamer.[141] Initially it was believed that a cysteine in the N-terminal portion of the polypeptide sequence was stabilising the ring though covalent disulfide bonding to adjacent dimers. However, even the mutation of this cysteine to alanine is not enough to diminish the toroid stability. [141] The influence of varying N-terminal residues on typical 2-Cys Prxs was perhaps most dramatically displayed by Gretes *et al.*[155] Small changes within this region on *Plasmoddium yoelii* Prx 1 (*Py*Prx1a) are enough to cause a "slip" at the B-type interface that leads to variations in the angle perpendicular to the dimer C_2-axis, at the inter-dimer interface. This is enough to change the geometry of the ring, with the tagged construct (*Py*Prx1a^{N*}) oligomerising as an octamer instead of a decamer.[155] While these effects are not seen in all Prxs [157] they are a common enough occurrence to warrant further study.

1.6.7 High molecular weight oligomers and other physiological roles

In addition to peroxidatic activity, Prxs have been shown to have a variety of different physiological roles. [129, 134, 145] This is perhaps unsurprising due to the apparent redundancy in the number of isologues found in many species. Humans, for example, have 6 different Prx isologues. The role that Prx adopts is intrinsically linked to its quaternary structure, with tubes, stacks and cages being noted in the literature. [109, 110, 158, 159] Only the physiological roles of typical 2-Cys from the Prx1 subfamily will be elaborated upon here.

Overoxidation of Cys_p

There are a number of studies that have described the formation of HMW oligomers *in vivo* and *in vitro* in response to a range of different conditions. One of the most well studied is the overoxidation of the Cys_p which can lead to toroid stabilisation and, in some cases, the formation of HMW structures. [133, 149, 160-162]

Prxs in prokaryotes tend to be more resistant to overoxidation than those found in eukaryotic cells. [145] A common sequence has been found between Prxs that are more sensitive to overoxidation, this being the YF motif in the C-term extension. This packs against α2 helix, stabilising the fully folded active site [163] and allowing the persistence of the conformation needed for an additional peroxide attack. [132] In fact, in a recent study by Cao *et al*. the Phe^{190} from the YF motif in *Bt*Prx3 was substituted for Leu to form a motif similar to those found in some bacterial Prxs that are more robust. [143] While the peroxidatic activity was not diminished, the sensitivity towards overoxidation was lowered, highlighting the importance of this motif with regards to overoxidation. The YF motif is highly conserved in many eukaryotic typical 2-Cys Prxs suggesting that the sensitivity to overoxidation has an important physiological role. [129, 143, 145] The overoxidised form can be reduced by sulfiredoxin (Srx). [129] Organisms that do not contain Srx have more robust Prxs [129] further suggesting that Cys_p-SO_2^- has a physiological significance and is not just an artefact of *in vitro* testing.

There are a number of hypotheses explaining why overoxidation is able to occur. [135] A number are related to stress and non-stress peroxide signalling. [160, 162] One such example of non-stress peroxide signalling is the "floodgate" mechanism whereby the inactivation of the Prx, due to localised H_2O_2 build up, leads to a further increase in H_2O_2 levels that are

utilised downstream. [162] The stress response is thought, in part, to be related to the ability of Prxs to act as a holdase chaperone, [150] preventing the aggregation of partially unfolded proteins under cell stress conditions. [157, 160, 164]

Chaperone

One of the many physiological roles of Prx is its ability to act as an ATP-independent holdase chaperone. [110, 145, 160, 164] For many years it has been suggested that the overoxidation of Prxs initiates the structural changes needed for chaperone holdase activity in both the LMW and HMW oligomers. [111, 150, 160, 165, 166] There are a number of examples where typical 2-Cys Prxs have been shown to have chaperone activity or at least have the capacity to behave as a putative chaperone.[131, 164] The chaperone activity is not directly linked to the peroxidatic activity, [164] with Cys_p and Cys_r non-active mutants both exhibiting the capacity to function as a chaperone. [157] However, the quaternary structure is important, with only the toroid or HMW oligomers displaying chaperone activity.[164] In this way, it can be described as a redox regulated chaperone as a disulfide bond between the active site cysteine residues (non-reduced) generally leads to disassembly of the toroid into its dimeric state. [148] The switch from peroxidatic function is not well understood, but it is believed to be in response to cellular stress for example, high temperature, [164] high [H_2O_2], [160] and changes in pH, [110] where Prxs switch between the dimer, LMW and HMW forms. [153] There is ongoing research into the structural differences between the two forms. [108, 111, 150] Understanding the structural changes may give insight not only into how the function diverges but also why

pH

Saccoccia *et al.* solved the crystal structure of a double stacked *Schistosoma mansoni* Prx1 (*Sm*Prx1) at pH 4.2. [110] The HMW oligomer is comprised of 20 Prx subunits arranged as a double stack of decameric rings. The rings are rotated 18 ^{o} with respect to each other around the 5-fold symmetry axis to form "cog-wheel" like interactions. The interactions at the ring interface (R-type interface) are predominantly polar, with the contacts between the carbonyl groups on α2 with the protonated His^{165} and Lys^{164} on α6'on the subunit directly above (fig. 1.18 between chains A and B).There are further polar contacts between Glu^{21} and Lys^{23} on the β2-strand in the bottom ring with the equivalent residues on β2' in the adjacent subunit in the top ring (fig. 1.18 between chains B and C).

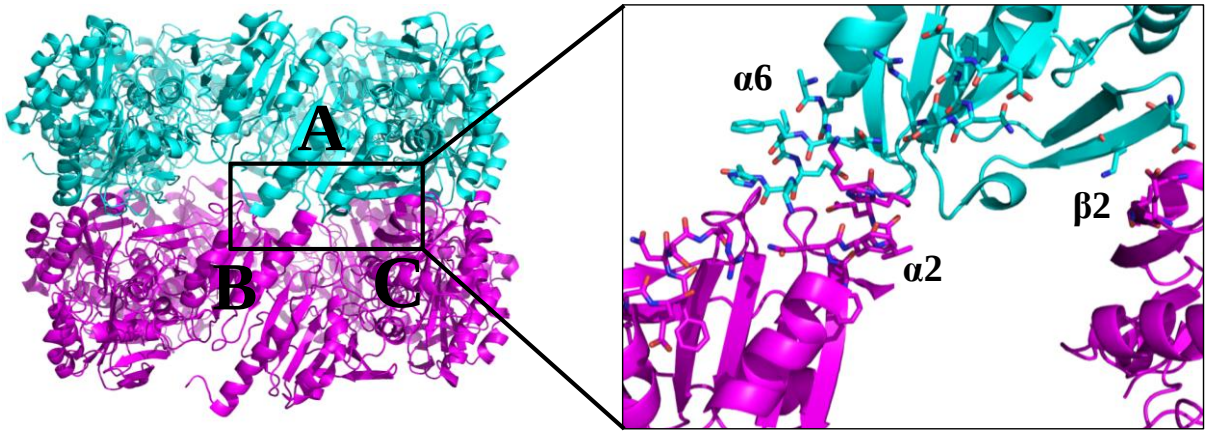


Figure 1.18: pH stack. Chain A has contacts with both chain B and chain C on the ring below. Only these three chains are shown for clarity in the enlarged image. His^{165} and Lys^{164} from α6 interact with the backbone carbonyl oxygen atoms on α2 (chain A and B). There are further polar contacts between Glu^{21} and Lys^{23} on the β2 (chain A and C).

By studying the differences in folding between the HMW and LMW crystal structures, they were able to identify possible changes that led to the stacking, namely the local unfolding of α2, the disordered C-terminus and the proximity of α6 to C-terminal Phe^{161}. These differences are specifically due to the acidic pH and more work will need to be done to compare this to the overoxidised structure.

Temperature

Mitochondrial tryparedoxin-peroxidase (mTXNPx) from *Leishmania infantum* has been shown to be involved in the defence of the parasite against changes in temperature (25 ^{o}C – 37 ^{o}C, change in temperature from insect to mammalian host) with the knockout strain producing a phenotype that has an increased sensitivity temperature. [164] *In vitro* studies of mTXNPx showed that the reduced decameric protein ($mTXNPx_{red}$) is able to decrease the degree of luciferase aggregation in response to an increase temperature while the oxidised, dimeric form ($mTXNPx_{ox}$) is not. Unlike other studies, [157, 160] the overoxidation of Cys_p was not needed to initiate chaperone activity, rather the increase in the surface hydrophobicity of $mTXNPx_{red}$ upon increasing temperature was believed to aid the switch to chaperone function. TEM imaging of $mTXNPx_{red}$ in the presence of luciferase showed a

decamer with electron density in the lumen that was not visible in the absence of the substrate protein suggesting that the holdase activity of this enzyme is derived from holding the partially heat denatured protein substrate within the toroid lumen.

1.6.8 Peroxiredoxins as tectons in nanotechnology

The evidence in the literature suggests that there are a number of routes towards forming HMW structures from Prxs. In fact, there have been a number of examples where this has occurred, through site-directed mutagenesis, changes in pH, the overoxidation of Cys_p, metal ion coordination to the N-terminal histidine tag and the introduction of crowding agents. [98, 109, 111, 159, 160] These structures range from rings and stacks, to catenanes, cages and tubes, all within the dimensions necessary to qualify as nanomaterials.

Catenanes

The first crystal structure of *Bt*Prx3 (93% sequence homology with *Hs*Prx3), published in 2005, showed that mutating Cys_r (C168) to a serine led to the formation of a previously unseen catenated conformation of dodecameric Prx toroids. [167] The points of contact between two basic dimeric units on alternating rings were identified as the residues that may drive HMW assembly. These are a mixture of H-bonded (Thr^{104} to $Lys^{88'}$ and Lys^{12} and Tyr^{10} to $Lys^{12'}$ and $Try^{10'}$) residues and a salt-bridge (Glu^{64} to $Arg^{109'}$). The catenated structure was seen again for *Bt*Prx3-F190L with similar points of contacts between the rings. [143] Furthermore, the mutation of C168S for *Hs*Prx3 also leads to a mixture of single and catenated rings as seem on TEM. [109] While the physiological relevance and exact mode of assembly remains unclear, mutations at the C-terminal region of this Prx do appear to be involved at least under crystallographic and TEM conditions. The HMW weight structures formed have a high surface area and a number of potential points for surface modification which could be utilised to form functional materials such as biosensors.

Cages

Prx decamers have been shown to assemble in dodecahedron cages under the conditions needed for TEM analysis. Meissner *et al.* used both crowing agents (polyethylene-glycol, PEG) and a TEM negative staining procedure that promotes 2D crystallisation (ammonium molybdate and trehalose) to form dodecahedral cages from human erythrocyte Prx2 (Prx-2).

[159] The structures formed can be varied by changing the TEM grid preparing. This includes adjusting the PEG concentration and the assembly time. For example, the samples with a low PEG concentration (0.05 %) only displayed single rings and some small clusters. At medium PEG (0.2 %) concentration and a short incubation time (~ 5 min) Prx-2 had partially assembled into cage-like structures but there was still a high proportion of single rings. Optimal conditions for cage formation were found to be 0.2 % PEG (w/v), 3 hour incubation time, 5 % (w/v) ammonium molybdate and 0.1 % (w/v) trehalose. The crystal structure of Prx-2 was docked into a 3D reconstruction of the cages to give a 16 Å resolution pseudo-atomic structure (fig. 1.19).

While the resolution is low it is clear that the cages are comprised of 12 Prx decamers and that the α6 helices are in close enough proximity to potentially aid cage stabilisation. The interior of the cage has the potential to be used to form size controlled nanoparticles.

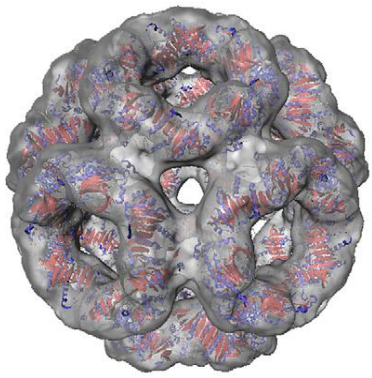


Figure 1.19: Prx cage. The crystal structure of Prx-2 was docked into a 3D reconstruction of the cages to give a 16 Å resolution pseudo-atomic structure of the cages formed from human erythrocyte Prx-2. [158] *Permission pending*

Tubes

Tubes from can be formed by utilising the N-terminal histidine tag as a ligand to coordinate to divalent metal ions. The crystal structure of *Sm*Prx1 has shown that the protein has the ability to stack into double decamers. [110] The oligomerisation can be amplified by increasing the R-type interface stability via coordination of divalent metal ions to the N-terminal histidine tag. [98] Aridni *el al* demonstrated that gold nanoparticles (AuNP), with

Ni^{2+} anchors, were able to coordinate to the histidine tags within minutes and tubes with conjugated AuNPs formed within hours. The tubes can be dissembled with addition of chelating agents such as imidazole or EDTA. [98]

As was described in section 1.4.4, *Hs*Prx3 assembles into elongated 1D nanotubes at acidic pH. This is believed to be due to the protonation of Cys_p and the subsequent local unfolding of the active site and C-terminal region. [109] A further study by Radjainia *et al.* elucidated the pseudo-atomic structure of these tubes to ~7 Å by mapping the crystal structure of *Bt*Prx3 onto the electron density map of the *Hs*Prx3 cryo-EM reconstruction (fig. 1.20). [108] By doing this they were able to confirm that the points of contact were indeed through the α2 and α6 helix plus additional contacts were noted between the partially resolved C-terminus and α2' and α3'/4'. This is different from the crystal structure of double stacked *Sm*Prx1 due to the increased ordering of the C-terminus, possibly a result of the incorporation of the YF motif in *Hs*Prx3 that is not seen in *Sm*Prx1. [108, 110, 143] They also noted significant differences in the ring diameter when compared to the *Bt*Prx3 crystal structure. These changes were attributed to changes at the A-type interfaces which appears to control the curvature of the toroid. Overall, the tubes show helical symmetry, with parameters 48.82 Å for the axial and azimuthal rotation angle of 8.06 ^{o}. Helical symmetry is open, meaning that the assembly can continue within the constraints of environmental conditions. [104]

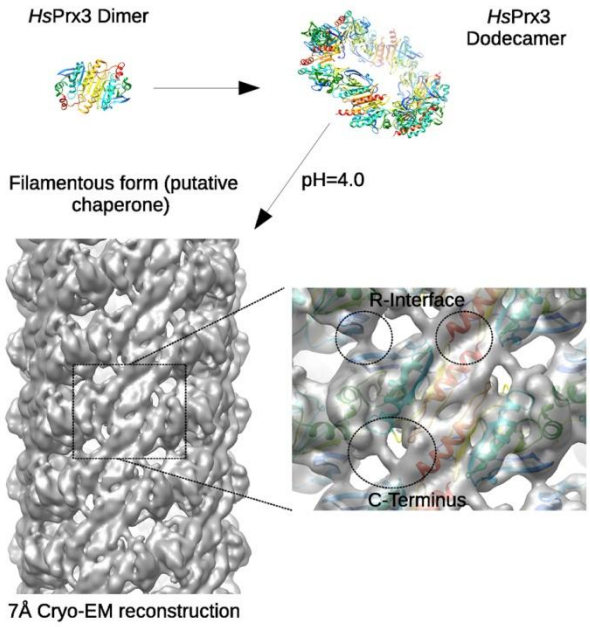


Figure 21:~7 Å ryo-EM reconstruction of a *Hs*Prx3 tube at pH 4.2. The crystal structure of *Bt*Prx3 is docked into the electron density map to show specific points of contacts between the rings. [108]

It is clear from the examples above that Prx is an ideal candidate for the formation of protein nanomaterials. By combining the pH sensitivity of *Hs*Prx3 with the metal chelating efficiency of the poly-6-histidine tag, further routes towards assembly will be explored.

1.7 Conclusion

The field of nanotechnology has grown immensely over the past 30 years. [5] Within this field there is a growing interest in the use of self-assembling biological molecules to form bespoke and functional nanomaterials. [37] Due to their 3D shape and the wealth of surface functionalities available, two proteins have been selected for study in this thesis to further progress the field of bionanotechnology by exploring new routes towards controlled self-assembly.

Lsr2 has been chosen due to its innate ability to self-assemble into long fibres and its DNA binding function. The N-terminal domain of this protein will also be examined in isolation to improve the enzymatic switch which aids assembly

*Hs*Prx3-6his has been identified as an ideal tecton for the formation of protein nanotubes. [109] This thesis strives to add to the previous work by discovering new routes towards controlled assembly and utilising the histidine tag as not only an oligomerisation trigger but also the potential source of functionalisation.

Overall, this work will describe how these proteins have been used to form novel nanomaterials, the routes by which these materials are formed and suggest possible functionalities that may be utilised in future work.

1.8 References

1. Miranda FF, Iwasaki K, Akashi S, Sumitomo K, Kobayashi M, Yamashita I, Tame JRH, Heddle JG: A self-assembled protein nanotube with high aspect ratio. *Small* 2009, 5(18):2077-2084.
2. Yang WR, Ren Q, Wu YN, Morris VK, Rey AA, Braet F, Kwan AH, Sunde M: Surface functionalization of carbon nanomaterials by self-assembling hydrophobin proteins. *Biopolymers* 2013, 99(1):84-94.
3. Allen M, Willits D, Young M, Douglas T: Constrained synthesis of cobalt oxide nanomaterials in the 12-subunit protein cage from Listeria innocua. *Inorganic Chemistry* 2003, 42(20):6300-6305.
4. Maghrebi M, Abbasi A, Amiri S, Monsefi R, Harati A: A collective and abridged lexical query for delineation of nanotechnology publications. *Scientometrics* 2011, 86(1):15-25.
5. Grieneisen ML, Zhang M: Nanoscience and nanotechnology: Evolving definitions and growing footprint on the scientific landscape. *Small* 2011, 7(20):2836-2839.
6. Behari J: Principles of nanoscience: An overview. *Indian Journal of Experimental Biology* 2010, 48(10):1008-1019.
7. Sanchez F, Sobolev K: Nanotechnology in concrete - A review. *Construction and Building Materials* 2010, 24(11):2060-2071.
8. Franks A: Nanotechnology. *Journal of Physics E-Scientific Instruments* 1987, 20(12):1442-1451.
9. Nicolau DE, Phillimore J, Cross R, Nicolau DV: Nanotechnology at the crossroads: the hard or the soft way? *Microelectronics Journal* 2000, 31(7):611-616.
10. Li X, Chen HC, Dang Y, Lin YL, Larson CA, Roco MC: A longitudinal analysis of nanotechnology literature: 1976-2004. *Journal of Nanoparticle Research* 2008, 10:3-22.
11. de Abril O, Gundel A, Maroun F, Allongue P, Schuster R: Single-step electrochemical nanolithography of metal thin films by localized etching with an AFM tip. *Nanotechnology* 2008, 19(32).
12. Balzani V: Nanoscience and nanotechnology: The bottom-up construction of molecular devices and machines. *Pure and Applied Chemistry* 2008, 80(8):1631-1650.
13. Whitesides GM, Grzybowski B: Self-assembly at all scales. *Science* 2002, 295(5564):2418-2421.

14. Chakrabarty R, Mukherjee PS, Stang PJ: Supramolecular coordination: Self-assembly of finite two- and three-dimensional ensembles. *Chemical Reviews* 2011, 111(11):6810-6918.
15. Lehn JM: Constitutional dynamic chemistry: Bridge from supramolecular chemistry to adaptive chemistry. *Constitutional Dynamic Chemistry* 2012, 322:1-32.
16. Safont-Sempere MM, Fernández G, Würthner F: Self-sorting phenomena in complex supramolecular systems. *Chemical Reviews* 2011, 111(9):5784-5814.
17. Dawe LN, Shuvaev KV, Thompson LK: Polytopic ligand directed self-assembly-polymetallic n x n grids versus non-grid oligomers. *Chemical Society Reviews* 2009, 38(8):2334-2359.
18. Kuppler RJ, Timmons DJ, Fang Q-R, Li J-R, Makal TA, Young MD, Yuan D, Zhao D, Zhuang W, Zhou H-C: Potential applications of metal-organic frameworks. *Coordination Chemistry Reviews* 2009, 253(23-24):3042-3066.
19. Janiak C, Vieth JK: MOFs, MILs and more: concepts, properties and applications for porous coordination networks (PCNs). *New Journal of Chemistry* 2010, 34(11):2366-2388.
20. Ramon Fortea-Perez F, Marino N, de Munno G, Armentano D, Julve M, Stiriba S-E: Intermolecular interactions in dictating the selfassembly of halogen derivatives of bis(N-substituted oxamato) palladate(II) complexes. *Rsc Advances* 2016, 6(8):6164-6170.
21. da Cunha TT, Oliveira WXC, Pinheiro CB, Pedroso EF, Nunes WC, Pereira CLM: Alkaline ion-modulated solid-state supramolecular organization in mixed organic/metallorganic compounds based on 1,1 '-ethylenebis(4-aminopyridinium) cations and bis(oxamate)cuprate(II) anions. *Crystal Growth & Design* 2016, 16(2):900-907.
22. Huard DJE, Kane KM, Tezcan FA: Re-engineering protein interfaces yields copper-inducible ferritin cage assembly. *Nature Chemical Biology* 2013, 9(3):169-176.
23. Heckl WM: Molecular self-assembly and nanomanipulation - Two key technologies in nanoscience and templating. *Advanced Engineering Materials* 2004, 6(10):843-847.
24. Mashaghi S, Jadidi T, Koenderink G, Mashaghi A: Lipid Nanotechnology. *International Journal of Molecular Sciences* 2013, 14(2):4242-4282.
25. Watson JD, Crick FHC: Genetical implications of the structure of deoxyribonucleic acid. *Nature* 1953, 171(4361):964-967.
26. Kool ET, Morales JC, Guckian KM: Mimicking the structure and function of DNA: Insights into DNA stability and replication. *Angewandte Chemie-International Edition* 2000, 39(6):990-1009.
27. Anfinsen CB: Formation and stabilisation of protein structure. *Biochemical Journal* 1972, 128(4):737-&.
28. Elemans J, Rowan AE, Nolte RJM: Mastering molecular matter. Supramolecular architectures by hierarchical self-assembly. *Journal of Materials Chemistry* 2003, 13(11):2661-2670.
29. Prockop DJ, Fertala A: The collagen fibril: The almost crystalline structure. *Journal of Structural Biology* 1998, 122(1-2):111-118.
30. Ottani V, Martini D, Franchi M, Ruggeri A, Raspanti M: Hierarchical structures in fibrillar collagens. *Micron* 2002, 33(7-8):587-596.
31. Ottani V, Raspanti M, Ruggeri A: Collagen structure and functional implications. *Micron* 2001, 32(3):251-260.
32. Beniash E: Biominerals-hierarchical nanocomposites: the example of bone. *Wiley Interdisciplinary Reviews-Nanomedicine and Nanobiotechnology* 2011, 3(1):47-69.

33. Ramakrishnan V, White SW: Ribosomal protein structures: insights into the architecture, machinery and evolution of the ribosome. *Trends in Biochemical Sciences* 1998, 23(6):208-212.
34. Nakamoto RK, Baylis Scanlon JA, Al-Shawi MK: The rotary mechanism of the ATP synthase. *Archives of Biochemistry and Biophysics* 2008, 476(1):43-50.
35. Corchero JL, Cedano J: Self-assembling, protein-based intracellular bacterial organelles: emerging vehicles for encapsulating, targeting and delivering therapeutical cargoes. *Microbial Cell Factories* 2011, 10.
36. Doles T, Bozic S, Gradisar H, Jerala R: Functional self-assembling polypeptide bionanomaterials. *Biochemical Society Transactions* 2012, 40:629 - 634.
37. Gradisar H, Jerala R: Self-assembled bionanostructures: proteins following the lead of DNA nanostructures. *Journal of Nanobiotechnology* 2014, 12(1):4.
38. Liu Q, Song C, Wang Z, Li N, Ding B: Precise organization of metal nanoparticles on DNA origami template. *Methods* 2013.
39. Morais MG, Martins VG, Steffens D, Pranke P, da Costa JAV: Biological applications of nanobiotechnology. *Journal of Nanoscience and Nanotechnology* 2014, 14(1):1007-1017.
40. Zayed JM, Nouvel N, Rauwald U, Scherman OA: Chemical complexity-supramolecular self-assembly of synthetic and biological building blocks in water. *Chemical Society Reviews* 2010, 39(8):2806-2816.
41. Yang L, Zhang LJ, Webster TJ: Nanobiomaterials: State of the art and future trends. *Advanced Engineering Materials* 2011, 13(6):B197-B217.
42. Scheibel T, Parthasarathy R, Sawicki G, Lin XM, Jaeger H, Lindquist SL: Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. *Proceedings of the National Academy of Sciences of the United States of America* 2003, 100(8):4527-4532.
43. Ke Y, Sharma J, Liu M, Jahn K, Liu Y, Yan H: Scaffolded DNA origami of a DNA tetrahedron molecular container. *Nano Letters* 2009, 9:2445 - 2447.
44. Goodman RP, Schaap IAT, Tardin CF, Erben CM, Berry RM, Schmidt CF, Turberfield AJ: Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. *Science* 2005, 310(5754):1661-1665.
45. Seeman NC: DNA in a material world. *Nature* 2003, 421(6921):427-431.
46. Shih W, Lin C: Knitting complex weaves with DNA origami. *Current Opinion in Structural Biology* 2010, 20:276 - 282.
47. Douglas S, Marblestone A, Teerapittayanon S, Vazquez A, Church G, Shih W: Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Research* 2009, 37:5001 - 5006.
48. Bell N, Engst C, Ablay M, Divitini G, Ducati C, Liedl T, Keyser U: DNA origami nanopores. *Nano Letters* 2012, 12:512 - 517.
49. Kuzuya A, Komiyama M: DNA origami: fold, stick, and beyond. *Nanoscale* 2010, 2:310 - 322.
50. Rothemund PWK: Folding DNA to create nanoscale shapes and patterns. *Nature* 2006, 440(7082):297-302.
51. Fu YM, Zeng DD, Chao J, Jin YQ, Zhang Z, Liu HJ, Li D, Ma HW, Huang Q, Gothelf KV *et al*: Single-step rapid assembly of DNA origami nanostructures for addressable nanoscale bioreactors. *Journal of the American Chemical Society* 2013, 135(2):696-702.
52. Grabow WW, Jaeger L: RNA self-assembly and RNA nanotechnology. *Accounts of Chemical Research* 2014, 47(6):1871-1880.

53. Grabow WW, Zakrevsky P, Afonin KA, Chworos A, Shapiro BA, Jaeger L: Self-assembling RNA nanorings based on RNAI/II inverse kissing complexes. *Nano Letters* 2011, 11(2):878-887.
54. Ohno H, Kobayashi T, Kabata R, Endo K, Iwasa T, Yoshimura SH, Takeyasu K, Inoue T, Saito H: Synthetic RNA-protein complex shaped like an equilateral triangle. *Nature Nanotechnology* 2011, 6(2):115-119.
55. Jaeger L, Chworos A: The architectonics of programmable RNA and DNA nanostructures. *Current Opinion in Structural Biology* 2006, 16(4):531-543.
56. Daldrop P, Lilley DMJ: The plasticity of a structural motif in RNA: Structural polymorphism of a kink turn as a function of its environment. *RNA-a Publication of the RNA Society* 2013, 19(3):357-364.
57. Lilley DMJ: The structure and folding of kink turns in RNA. *Wiley Interdisciplinary Reviews-RNA* 2012, 3(6):797-805.
58. Martinek TA, Hetenyi A, Fulop L, Mandity IM, Toth GK, Dekany I, Fulop F: Secondary structure dependent self-assembly of beta-peptides into nanosized fibrils and membranes. *Angewandte Chemie-International Edition* 2006, 45(15):2396-2400.
59. Aggeli A, Nyrkova IA, Bell M, Harding R, Carrick L, McLeish TCB, Semenov AN, Boden N: Hierarchical self-assembly of chiral rod-like molecules as a model for peptide β-sheet tapes, ribbons, fibrils, and fibers. *Proceedings of the National Academy of Sciences of the United States of America* 2001, 98(21):11857-11862.
60. Scanlon S, Aggeli A, Boden N, McLeish TCB, Hine P, Koopmans RJ, Crowder C: Organisation of self-assembling peptide nanostructures into macroscopically ordered lamella-like layers by ice crystallisation. *Soft Matter* 2009, 5(6):1237-1246.
61. Xu G, Pranantyo D, Zhang B, Xu LQ, Neoh KG, Kang ET: Tannic acid anchored layer-by-layer covalent deposition of parasin I peptide for antifouling and antimicrobial coatings. *Royal Society of Chemistry Advances* 2016, 6(18):14809-14818.
62. Eakins GL, Pandey R, Wojciechowski JP, Zheng HY, Webb JEA, Valery C, Thordarson P, Plank NOV, Gerrard JA, Hodgkiss JM: Functional organic semiconductors assembled via natural aggregating peptides. *Advanced Functional Materials* 2015, 25(35):5640-5649.
63. Sedman VL, Chen X, Allen S, Roberts CJ, Korolkov VV, Tendler SJB: Tuning the mechanical properties of self-assembled mixed-peptide tubes. *Journal of Microscopy* 2013, 249(3):165-172.
64. Ghadiri MR, Granja JR, Milligan RA, McRee DE, Khazanovich N: Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* 1993, 366(6453):324-327.
65. Ghadiri M, Granja J, Buehler L: Artificial transmembrane ion channels from self-assembling peptide nanotubes. *Nature* 1994, 369:301 - 304.
66. Clark TD, Buehler LK, Ghadiri MR: Self-assembling cyclic beta(3)-peptide nanotubes as artificial transmembrane ion channels. *Journal of the American Chemical Society* 1998, 120(4):651-656.
67. Hourani R, Zhang C, van der Weegen R, Ruiz L, Li CY, Keten S, Helms BA, Xu T: Processable cyclic peptide nanotubes with tunable interiors. *Journal of the American Chemical Society* 2011, 133(39):15296-15299.
68. Qu Y, Lim CJ, Whang YR, Liu J, Yan J: Mechanism of DNA organization by Mycobacterium tuberculosis protein Lsr2. *Nucleic Acids Research* 2013, 41(10):5263-5272.

69. Sarkar B, O'Leary LER, Hartgerink JD: Self-assembly of fiber-forming collagen mimetic peptides controlled by triple-helical nucleation. *Journal of the American Chemical Society* 2014, 136(41):14417-14424.
70. Valéry C, Artzner F, Robert B, Gulick T, Keller G, Grabielle-Madelmont C, Torres ML, Cherif-Cheikh R, Paternostre M: Self-association process of a peptide in solution: From β-sheet filaments to large embedded nanotubes. *Biophysical Journal* 2004, 86(4):2484-2501.
71. Dobson CM: Protein folding and misfolding. *Nature* 2003, 426(6968):884-890.
72. O'Leary LER, Fallas JA, Bakota EL, Kang MK, Hartgerink JD: Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel. *Nature Chemistry* 2011, 3(10):821-828.
73. Valery C, Pandey R, Gerrard JA: Protein beta-interfaces as a generic source of native peptide tectons. *Chemical Communications* 2013, 49(27):2825-2827.
74. Howorka S: Rationally engineering natural protein assemblies in nanobiotechnology. *Current Opinion in Biotechnology* 2011, 22(4):485-491.
75. Summers EL, Meindl K, Uson I, Mitra AK, Radjainia M, Colangeli R, Alland D, Arcus VL: The structure of the oligomerization domain of Lsr2 from Mycobacterium tuberculosis reveals a mechanism for chromosome organization and protection. *PloS ONE* 2012, 7(6).
76. Doyle DA, Cabral JM, Pfuetzner RA, Kuo AL, Gulbis JM, Cohen SL, Chait BT, MacKinnon R: The structure of the potassium channel: Molecular basis of K+ conduction and selectivity. *Science* 1998, 280(5360):69-77.
77. Wood ZA, Schroder E, Harris JR, Poole LB: Structure, mechanism and regulation of peroxiredoxins. *Trends in Biochemical Sciences* 2003, 28(1):32-40.
78. Cingolani R, Rinaldi R, Maruccio G, Biasco A: Nanotechnology approaches to self-organized bio-molecular devices. *Physica E-Low-Dimensional Systems & Nanostructures* 2002, 13(2-4):1229-1235.
79. Kortemme T, Baker D: Computational design of protein–protein interactions. *Current Opinion in Chemical Biology* 2004, 8(1):91-97.
80. Ross PD, Subramanian S: Thermodynamics of protein association reactions: forces contributing to stability. *Biochemistry* 1981, 20(11):3096-3102.
81. Tsai CJ, Zheng J, Zanuy D, Haspel N, Wolfson H, Aleman C, Nussinov R: Principles of nanostructure design with protein building blocks. *Proteins-Structure Function and Bioinformatics* 2007, 68(1):1-12.
82. Rose GD, Fleming PJ, Banavar JR, Maritan A: A backbone-based theory of protein folding. *Proceedings of the National Academy of Sciences of the United States of America* 2006, 103(45):16623-16633.
83. Brodin JD, Smith SJ, Carr JR, Tezcan FA: Designed, helical protein nanotubes with variable diameters from a single building block. *Journal of the American Chemical Society* 2015, 137(33):10468-10471.
84. Buch I, Brooks BR, Wolfson HJ, Nussinov R: Computational validation of protein nanotubes. *Nano Letters* 2009, 9(3):1096-1102.
85. Dill K, MacCallum J: The protein-folding problem, 50 years on. *Science* 2012, 338:1042 - 1046.
86. Apostolovic B, Danial M, Klok H: Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials. *Chemical Society Reviews* 2010, 39:3541 - 3575.
87. de Vega MJP, Martin-Martinez M, Gonzalez-Muniz R: Modulation of protein-protein interactions by stabilizing/mimicking protein secondary structure elements. *Current Topics in Medicinal Chemistry* 2007, 7(1):33-62.

88. Nesloney CL, Kelly JW: Progress towards understanding beta-sheet structure. *Bioorganic & Medicinal Chemistry* 1996, 4(6):739-766.
89. Greenwald J, Riek R: On the possible amyloid origin of protein folds. *Journal of Molecular Biology* 2012, 421(4-5):417-426.
90. Marianayagam NJ, Sunde M, Matthews JM: The power of two: protein dimerization in biology. *Trends in Biochemical Sciences* 2004, 29(11):618-625.
91. Papapostolou D, Howorka S: Engineering and exploiting protein assemblies in synthetic biology. *Molecular Biosystems* 2009, 5(7):723-732.
92. Ariga K, Ito H, Hill JP, Tsukube H: Molecular recognition: from solution science to nano/materials technology. *Chemical Society Reviews* 2012, 41(17):5800-5835.
93. Chen J, Sawyer N, Regan L: Proteinprotein interactions: General trends in the relationship between binding affinity and interfacial buried surface area. *Protein Science* 2013, 22(4):510-515.
94. DeLano WL: Unraveling hot spots in binding interfaces: Progress and challenges. *Current Opinion in Structural Biology* 2002, 12(1):14-20.
95. Ballister ER, Lai AH, Zuckermann RN, Cheng Y, Mougous JD: In vitro self-assembly from a simple protein of tailorable nanotubes building block. *Proceedings of the National Academy of Sciences of the United States of America* 2008, 105(10):3733-3738.
96. Brodin JD, Ambroggio XI, Tang C, Parent KN, Baker TS, Tezcan FA: Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* 2012, 4(5):375-382.
97. Brodin JD, Carr JR, Sontz PA, Tezcan FA: Exceptionally stable, redox-active supramolecular protein assemblies with emergent properties. *Proceedings of the National Academy of Sciences of the United States of America* 2014, 111(8):2897-2902.
98. Ardini M, Giansanti F, Di Leandro L, Pitari G, Cimini A, Ottaviano L, Donarelli M, Santucci S, Angelucci F, Ippoliti R: Metal-induced self-assembly of Peroxiredoxin as a tool for sorting ultrasmall gold nanoparticles into one-dimensional clusters. *Nanoscale* 2014, 6(14):8052-8061.
99. Zhang W, Luo Q, Miao L, Hou C, Bai Y, Dong Z, Xu J, Liu J: Self-assembly of glutathione S-transferase into nanowires. *Nanoscale* 2012, 4(19):5847-5851.
100. Svanedal I, Boija S, Almesaker A, Persson G, Andersson F, Hedenstrom E, Bylund D, Norgren M, Edlund H: Metal ion coordination, conditional stability constants, and solution behavior of chelating surfactant metal complexes. *Langmuir* 2014, 30(16):4605-4612.
101. Bai YS, Luo Q, Zhang W, Miao L, Xu JY, Li HB, Liu JQ: Highly ordered protein nanorings designed by accurate control of Glutathione S-Transferase self-assembly. *Journal of the American Chemical Society* 2013, 135(30):10966-10969.
102. Watanabe H, Matsumaru H, Ooishi A, Feng YW, Odahara T, Suto K, Honda S: Optimizing pH Response of Affinity between Protein G and IgG Fc. How electrostatic moddulation affect protein-protein interactions. *Journal of Biological Chemistry* 2009, 284(18):12373-12383.
103. Morais MAB, Giuseppe PO, Souza TACB, Alegria TGP, Oliveira MA, Netto LES, Murakami MT: How pH modulates the dimer-decamer interconversion of 2-Cys Peroxiredoxins from the Prx1 subfamily. *Journal of Biological Chemistry* 2015, 290(13):8582-8590.
104. Lai YT, King NP, Yeates TO: Principles for designing ordered protein assemblies. *Trends in Cell Biology* 2012, 22(12):653-661.

105. Padilla JE, Colovos C, Yeates TO: Nanohedra: Using symmetry to design self assembling protein cages, layers, crystals, and filaments. *Proceedings of the National Academy of Sciences of the United States of America* 2001, 98(5):2217-2221.
106. Sasso L, Suei S, Domigan L, Healy J, Nock V, Williams MAK, Gerrard JA: Versatile multi-functionalization of protein nanofibrils for biosensor applications. *Nanoscale* 2014, 6(3):1629-1634.
107. Angelucci F, Bellelli A, Ardini M, Ippoliti R, Saccoccia F, Morea V: One ring (or two) to hold them all - on the structure and function of protein nanotubes. *The Federation of European Biochemical Societies Journal* 2015, 282(15):2827-2845.
108. Radjainia M, Venugopal H, Desfosses A, Phillips AJ, Yewdall NA, Hampton MB, Gerrard JA, Mitra AK: Cryo-electron microscopy structure of human Peroxiredoxin-3 filament reveals the assembly of a putative chaperone. *Structure* 2015, 23(5):912-920.
109. Phillips AJ, Littlejohn J, Yewdall NA, Zhu T, Valery C, Pearce FG, Mitra AK, Radjainia M, Gerrard JA: Peroxiredoxin is a versatile self-assembling tecton for protein nanotechnology. *Biomacromolecules* 2014, 15(5):1871-1881.
110. Saccoccia F, Di Micco P, Boumis G, Brunori M, Koutris I, Miele AE, Morea V, Sriratana P, Williams DL, Bellelli A *et al*: Moonlighting by different stressors: Crystal structure of the chaperone species of a 2-Cys Peroxiredoxin. *Structure* 2012, 20(3):429-439.
111. Angelucci F, Saccoccia F, Ardini M, Boumis G, Brunori M, Di Leandro L, Ippolitl R, Miele AE, Natoli G, Scotti S *et al*: Switching between the alternative structures and functions of a 2-Cys Peroxiredoxin, by site-directed mutagenesis. *Journal of Molecular Biology* 2013, 425(22):4556-4568.
112. Gordon BRG, Imperial R, Wang L, Navarre WW, Liu J: Lsr2 of Mycobacterium represents a novel class of H-NS-like proteins. *Journal of Bacteriology* 2008, 190(21):7052-7059.
113. Gordon BRG, Li YF, Wang LR, Sintsova A, van Bakel H, Tian SH, Navarre WW, Xia B, Liu J: Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in Mycobacterium tuberculosis *Proceedings of the National Academy of Sciences of the United States of America* 2010, 107(43):18741-18741.
114. Gordon BRG, Li YF, Cote A, Weirauch MT, Ding PF, Hughes TR, Navarre WW, Xia B, Liu J: Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proceedings of the National Academy of Sciences of the United States of America* 2011, 108(26):10690-10695.
115. Du YL, Zhang H, He Y, Huang F, He ZG: Mycobacterium smegmatis Lsr2 physically and functionally interacts with a new flavoprotein involved in bacterial resistance to oxidative stress. *Journal of Biochemistry* 2012, 152(5):479-486.
116. Colangeli R, Haq A, Arcus VL, Summers E, Magliozzo RS, McBride A, Mitra AK, Radjainia M, Khajo A, Jacobs WR *et al*: The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proceedings of the National Academy of Sciences of the United States of America* 2009, 106(11):4414-4418.
117. Saini C, Prasad HK, Rani R, Murtaza A, Misra N, Narayan NPS, Nath I: Lsr2 of Mycobacterium leprae and Its synthetic peptides elicit restitution of T cell responses in Erythema nodosum leprosum and reversal reactions in patients with Lepromatous leprosy. *Clinical and Vaccine Immunology* 2013, 20(5):673-682.
118. Colangeli R, Helb D, Vilcheze C, Hazbon MH, Lee CG, Safi H, Sayers B, Sardone I, Jones MB, Fleischmann RD *et al*: Transcriptional regulation of multi-drug tolerance and antibiotic-induced responses by the histone-like protein Lsr2 in M-tuberculosis. *PLoS Pathogens* 2007, 3(6):780-793.

119. Ali SS, Xia B, Liu J, Navarre WW: Silencing of foreign DNA in bacteria. *Current Opinion in Microbiology* 2012, 15(2):175-181.
120. Browning DF, Grainger DC, Busby SJW: Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Current Opinion in Microbiology* 2010, 13(6):773-780.
121. Shintani M, Suzuki-Minakuchi C, Nojiri H: Nucleoid-associated proteins encoded on plasmids: Occurrence and mode of function. *Plasmid* 2015, 80:32-44.
122. Dorman CJ: H-NS-like nucleoid-associated proteins, mobile genetic elements and horizontal gene transfer in bacteria. *Plasmid* 2014, 75:1-11.
123. Lawrence JG, Ochman H: Molecular archaeology of the Escherichia coli genome. *Proceedings of the National Academy of Sciences of the United States of America* 1998, 95(16):9413-9417.
124. Chen JM, Ren H, Shaw JE, Wang YJ, Li M, Leung AS, Tran V, Berbenetz NM, Kocincova D, Yip CM *et al*: Lsr2 of Mycobacterium tuberculosis is a DNA-bridging protein. *Nucleic Acids Research* 2008, 36(7):2123-2135.
125. Bartek IL, Woolhiser LK, Baughn AD, Basaraba RJ, Jacobs WR, Jr., Lenaerts AJ, Voskuil MI: Mycobacterium tuberculosis Lsr2 is a global transcriptional regulator required for adaptation to changing oxygen levels and virulence. *mBio* 2014, 5(3).
126. Chen JM, German GJ, Alexander DC, Ren H, Tan T, Liu J: Roles of Lsr2 in colony morphology and biofilm formation of Mycobacterium smegmatis. *Journal of Bacteriology* 2006, 188(2):633-641.
127. Ashmead HM, Negron L, Webster K, Arcus V, Gerrard JA: Proteins as supramolecular building blocks: Nterm-Lsr2 as a new protein tecton. *Biopolymers* 2015, 103(5):260-270.
128. Karplus PA: A primer on peroxiredoxin biochemistry. *Free Radical Biology and Medicine* 2015, 80:183-190.
129. Perkins A, Poole LB, Karplus PA: Tuning of peroxiredoxin catalysis for various physiological roles. *Biochemistry* 2014, 53(49):7693-7705.
130. Wood ZA, Poole LB, Karplus PA: Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science* 2003, 300(5619):650-653.
131. Hall A, Nelson K, Poole LB, Karplus PA: Structure-based insights into the catalytic power and conformational dexterity of Peroxiredoxins. *Antioxidants & Redox Signaling* 2011, 15(3):795-815.
132. Hall A, Parsonage D, Poole LB, Karplus PA: Structural evidence that Peroxiredoxin catalytic power is based on transition-state stabilization. *Journal of Molecular Biology* 2010, 402(1):194-209.
133. Karplus PA, Hall A: Structural survey of the peroxiredoxins. *Sub-Cellular Biochemistry* 2007, 44:41-60.
134. Perkins A, Nelson KJ, Parsonage D, Poole LB, Karplus PA: Peroxiredoxins: guardians against oxidative stress and modulators of peroxide signaling. *Trends in Biochemical Sciences* 2015, 40(8):435-445.
135. Sharapov MG, Ravin VK, Novoselov VI: Peroxiredoxins as multifunctional enzymes. *Molecular Biology* 2014, 48(4):520-545.
136. Leslie BP, and Kimberly JN: Distribution and features of the six classes of Peroxiredoxins. *Molecular Cell* 2016, 39(1):53-59.
137. Nelson KJ, Knutson ST, Soito L, Klomsiri C, Poole LB, Fetrow JS: Analysis of the peroxiredoxin family: Using active-site structure and sequence information for global classification and residue analysis. *Proteins-Structure Function and Bioinformatics* 2011, 79(3):947-964.

138. Cox AG, Peskin AV, Paton LN, Winterbourn CC, Hampton MB: Redox potential and Peroxide reactivity of human Peroxiredoxin 3. *Biochemistry* 2009, 48(27):6495-6501.
139. Alphey MS, Bond CS, Tetaud E, Fairlamb AH, Hunter WN: The structure of reduced tryparedoxin peroxidase reveals a decamer and insight into reactivity of 2Cys-peroxiredoxins. *Journal of Molecular Biology* 2000, 300(4):903-916.
140. Copley SD, Novak WRP, Babbitt PC: Divergence of function in the thioredoxin fold suprafamily: Evidence for evolution of peroxiredoxins from a thioredoxin-like ancestor. *Biochemistry* 2004, 43(44):13981-13995.
141. Cao Z, Tavender TJ, Roszak AW, Cogdell RJ, Bulleid NJ: Crystals structure of reduced and of oxidized Peroxiredoxin IV enzyme reveals a stable oxidized decamer and a non-disulfide-bonded intermediate in the catalytic cycle. *Journal of Biological Chemistry* 2011, 286(49):42257-42266.
142. Schroder E, Littlechild JA, Lebedev AA, Errington N, Vagin AA, Isupov MN: Crystal structure of decameric 2-Cys peroxiredoxin from human erythrocytes at 1.7 angstrom resolution. *Structure with Folding & Design* 2000, 8(6):605-615.
143. Cao Z, McGow DP, Shepherd C, Lindsay JG: Improved catenated structures of Bovine Peroxiredoxin III F190L reveal details of ring-ring interactions and a novel conformational state. *PLOS One* 2015, 10(4).
144. Hall A, Sankaran B, Poole LB, Karplus PA: Structural changes common to catalysis in the Tpx Peroxiredoxin subfamily. *Journal of Molecular Biology* 2009, 393(4):867-881.
145. Hall A, Karplus PA, Poole LB: Typical 2-Cys peroxiredoxins - structures, mechanisms and functions. *Federation of European Biochemical Societies Journal* 2009, 276(9):2469-2477.
146. Perkins A, Nelson KJ, Williams JR, Parsonage D, Poole LB, Karplus PA: The sensitive balance between the fully folded and locally unfolded conformations of a model Peroxiredoxin. *Biochemistry* 2013, 52(48):8708-8721.
147. Sarma GN, Nickel C, Rahlfs S, Fischer M, Becker K, Karplus PA: Crystal structure of a novel Plasmodium falciparum 1-Cys Peroxiredoxin. *Journal of Molecular Biology* 2005, 346(4):1021-1034.
148. Wood ZA, Poole LB, Hantgan RR, Karplus PA: Dimers to doughnuts: Redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry* 2002, 41(17):5493-5504.
149. Hugo M, Turell L, Manta B, Botti H, Monteiro G, Netto LES, Alvarez B, Radi R, Trujillo M: Thiol and sulfenic acid oxidation of AhpE, the one-cysteine Peroxiredoxin from Mycobacterium tuberculosis: Kinetics, acidity constants, and conformational Dynamics. *Biochemistry* 2009, 48(40):9416-9426.
150. Konig J, Galliardt H, Jutte P, Schaper S, Dittmann L, Dietz KJ: The conformational bases for the two functionalities of 2-cysteine peroxiredoxins as peroxidase and chaperone. *Journal of Experimental Botany* 2013, 64(11):3483-3497.
151. Cao Z, Bhella D, Lindsay JG: Reconstitution of the mitochondrial PrxIII antioxidant defence pathway: General properties and factors affecting PrxII activity and oligomeric state. *Journal of Molecular Biology* 2007, 372(4):1022-1033.
152. Matsumura T, Okamoto K, Iwahara S-I, Hori H, Takahashi Y, Nishino T, Abe Y: Dimer-oligomer interconversion of wild-type and mutant rat 2-Cys peroxiredoxin. *Journal of Biological Chemistry* 2008, 283(1):284-293.
153. Barranco-Medina S, Lazaro J-J, Dietz K-J: The oligomeric conformation of peroxiredoxins links redox state to function. *FEBS Letters* 2009, 583(12):1809-1816.

154. Noichri Y, Palais G, Ruby V, D'Autreaux B, Delaunay-Moisan A, Nystrom T, Molin M, Toledano MB: In vivo parameters influencing 2-Cys Prx oligomerization: The role of enzyme sulfinylation. *Redox Biology* 2015, 6:326-333.
155. Gretes MC, Karplus PA: Observed octameric assembly of a Plasmodium yoelii peroxiredoxin can be explained by the replacement of native "ball-and-socket" interacting residues by an affinity tag. *Protein Science* 2013, 22(10):1445-1452.
156. Gourlay LJ, Bhella D, Kelly SM, Price NC, Lindsay JG: Structure-function analysis of recombinant substrate protein 22 kDa (SP-22) - A mitochondrial 2-Cys peroxiredoxin organized as a decameric toroid. *Journal of Biological Chemistry* 2003, 278(35):32631-32637.
157. Banerjee M, Chakravarty D, Ballal A: Redox-dependent chaperone/peroxidase function of 2-Cys-Prx from the cyanobacterium Anabaena PCC7120: role in oxidative stress tolerance. *BioMed Central plant biology* 2015, 15(1):444-444.
158. Phillips AJ, Littlejohn J, Yewdall NA, Zhu T, Valery C, Pearce FG, Mitra AK, Radjainia M, Gerrard JA: Peroxiredoxin is a sersatile self-assembling tecton for protein nanotechnology. *Biomacromolecules* 2014, 15(5):1871-1881.
159. Meissner U, Schroder E, Scheffler D, Martin AG, Harris JR: Formation, TEM study and 3D reconstruction of the human erythrocyte peroxiredoxin-2 dodecahedral higher-order assembly. *Micron* 2007, 38(1):29-39.
160. Jang HH, Lee KO, Chi YH, Jung BG, Park SK, Park JH, Lee JR, Lee SS, Moon JC, Yun JW *et al*: Two enzymes in one: Two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell* 2004, 117(5):625-635.
161. Hall A, Parsonage D, Horita D, Karplus PA, Poole LB, Barbar E: Redox-dependent dynamics of a dual thioredoxin fold protein: Evolution of specialized folds. *Biochemistry* 2009, 48(25):5984-5993.
162. Kil IS, Lee SK, Ryu KW, Woo HA, Hu M-C, Bae SH, Rhee SG: Feedback control of adrenal Steroidogenesis via H2O2-dependent, reversible inactivation of Peroxiredoxin III in mitochondria. *Molecular Cell* 2012, 46(5):584-594.
163. Wang X, Wang LK, Wang X, Sun F, Wang CC: Structural insights into the peroxidase activity and inactivation of human peroxiredoxin 4. *Biochemical Journal* 2012, 441(1):113-118.
164. Teixeira F, Castro H, Cruz T, Tse E, Koldewey P, Southworth DR, Tomas AM, Jakob U: Mitochondrial peroxiredoxin functions as crucial chaperone reservoir in Leishmania infantum. *Proceedings of the National Academy of Sciences of the United States of America* 2015, 112(7):E616-E624.
165. Saccoccia F, Angelucci F, Boumis G, Desiato G, Miele AE, Bellelli A: Selenocysteine robustness versus cysteine versatility: a hypothesis on the evolution of the moonlighting behaviour of peroxiredoxins. *Biochemical Society Transactions* 2014, 42(6):1768-1772.
166. Toledano MB, Huang B: Microbial 2-Cys peroxiredoxins: Insights into their complex physiological roles. *Molecules and Cells* 2016, 39(1):31-39.
167. Cao ZB, Roszak AW, Gourlay LJ, Lindsay JG, Isaacs NW: Bovine mitochondrial Peroxiredoxin III forms a two-ring catenane. *Structure* 2005, 13(11):1661-1664.

Chapter Two: Materials and methods

The following chapter describes the methods and materials used throughout this study.

2.1 Materials

2.1.1 Chemical reagents

All chemical reagents were purchased from Sigma Aldrich (St. Louis, MO) unless otherwise specified. Tryptone media was purchased from Oxoid (United Kingdom) and yeast extract was purchased from Bacto-BD (United States). Plasmids were synthesised by Epoch Lifescience Inc (Sugar Land, TX) or GenScript (Piscataway, NJ).

2.1.2 Bacterial strains

BL21(DE3)*

BL21(pRARE)Rosetta

Competent bacterial strains were prepared in-house and stored in 25% glycerol at – 80 ^{o}C.

2.1.3 Antibiotics

Each plasmid was engineered to have a particular antibiotic resistance (Table 2.1). Solutions of antibiotics were prepared at the following concentrations:

Antibiotic	Abbreviation	Solvent	Stock conc.	Working conc.
Ampicillin	Amp	H_2O	100 mg/mL	100 μg/mL
Kanamycin	Kan	H_2O	50 mg/mL	50 μg/mL
Chloramphenicol	Cam	EtOH	35 mg/mL	35 μg/mL

Table 2.1: List of antibiotics used during protein expression (AppliChem, Darmstadt, Germany). Growth solutions and agar plates were supplemented with the antibiotic relating to the plasmid resistance.

2.1.4 Plasmids

Type	Plasmid	Resistance	Chapter
Nterm-Lsr2			
	pET30b – Lsr2	Kan	
	pProExHtb-LNterm	Amp	3
	pProExHt-Nterm2	Amp	
***Hs*Prx3 WT – 6 his**			
	pET151 D-TOPO – *Hs*Prx3 Nterm	Amp + Cam	4 + 5
	pET11a-hsprx3 Cterm	Amp + Cam	
His mutants			5
	pET28a-PRX3E	Kan + Cam	
	pET28a-PRX3F	Kan + Cam	
R-interface mutants			
	pET151 D-TOPO-hsPrx3 K23A	Amp + Cam	
	pET151 D-TOPO-hsPrx3 K23H	Amp + Cam	4
	pET151 D-TOPO-hsPrx3 K23R	Amp + Cam	
Linker mutants			
	pET28a-PRX3A	Kan + Cam	
	pET28a-PRX3B	Kan + Cam	
	pET28a-PRX3C	Kan + Cam	5
	pET28a-PRX3D	Kan + Cam	
	pET28a-PRX3D(2)	Kan + Cam	
recombinant TEV	pRK796238Δ	Kan + Cam	4

Table 2.2: List of the plasmids used in this thesis. Lsr2 and Nterm-Lsr2 were synthesised in-house at the University of Waikato, Ent-Nterm was synthesised by GenScript (Piscataway, NJ), R-interface mutants were developed in-house from the commercial *Hs*Prx3-WT plasmid at the University of Auckland and all other peroxiredoxin plasmids were synthesised by Epoch Lifescience (Sugar Land, TX).

2.1.5 Enzymes

2.1.5.1 Proteases

Trypsin
Enteropeptidase **(**Applied Biological Materials (ABM) Inc, Canada)
Tobacco Etch Virus (TEV) protease (in-house)

2.1.5.2 Restriction enzymes

NdeI – digest in buffer H
NcoI – digest in buffer K
XhoI – digest in buffer H

For double digests a universal buffer was used
All enzymes (10 U/µL) and buffers purchased from Invitrogen (Carlbad, US)

2.1.5.3 Peroxiredoxin peroxidise assay

Horse Radish Peroxidase (HRP)

2.1.6 Media

All media and media containers were sterilised by high pressure saturated steam (121 ^{o}C) for 20 minutes prior to use. Media could be stored at 4 ^{o}C for up to two days if not used immediately.

2.1.6.1 Lysogeny broth (LB)

Lysogeny broth was used to grow all *E. coli* batches unless otherwise stated. The media contained tryptone (Oxoid, United Kindom), yeast extract (Bacto-BD, United States) and sodium chloride at 1%, 0.5% and 1% respectively in deionised water.

2.1.6.2 Lysogeny broth agar (LA)

The media was prepared as above with the addition of 1.5% agar (Bacto-BD, United States). This was dissolved into the media through the autoclave process. Appropriate antibiotics were added to the mix once the solution had cooled to 45^{o}C and the plates were poured under sterile conditions [1] with ~ 20 mL per plate.

2.1.6.3 Auto-induction media

Auto-induction provided a means to grow bacteria without the need for manual induction. [2] The media consisted of 10 g of Tryptone and 5 g of yeast extract, dissolved in 960 mL MilliQ water. The media was autoclaved before the addition of 20 mL 50 x 5052 stock (25 % glycerol, 2.5 % glucose and 10% lactose), 20 mL 50 x M stock (1.25 M Na_2HPO_4, 1.25 M KH_2PO_4, 2.5 M NH_4Cl and 0.25 M $NaSO_4$), 1 mL 2 M $MgSO_4$, 200 µL trace metals (50 mM $FeCl_2$, 20 mM $CaCl_2$, 10 mM each of $MnCl_2$ and $ZnSO_4$, 2 mM each of $CoCl_2$, $CuCl_2$, $NiCl_2$, $NaMoO_4$, Na2SeO_3, H3BO_3 in 60 mM HCl) directly before inoculation.

2.1.7 Buffers

All buffers were made using the reagents listed in the table, with the weight measured to ± 0.01 g. The samples were incubated at the appropriate temperature before the pH was analysed. At this point the buffers were vacuum filtered through 0.2 μM filter paper.

2.1.7.1 Lsr2, Nterm-Lsr2 and Ent-Nterm buffers

Buffer	Composition
Lysis	20 mM imidazole, 200 mM NaCl, 50 mM NaP_i, pH 7.4
IMAC elution	1 M imidazole, 200 mM NaCl, 50 mM NaP_i, pH 7.4
Storage/proteolysis	200 mM NaCl, 50 mM NaP_i, pH 7.4
High salt	1 M NaCl, 50 mM NaP_i, pH 7.4
pH 9.0	200 mM NaCl, 200 mM Tris base
pH 8.0	200 mM NaCl, 200 mM Tris base
pH 6.0	200 mM NaCl, 200 mM NaP_i
CD	20 mM NaP_i

Table 2.3 *NaP_i relates to the ratio of NaH_2PO_4 and Na_2HPO_4 that has been calculated to achieve the correct pH. [3]*

2.1.7.2 Buffers for Prx purification (all at pH 8.0)

Buffer	Composition
Lysis	20 mM imidazole, 150 mM NaCl, 20 mM HEPES
IMAC elution	500 mM imidazole, 150 mM NaCl, 20 mM HEPES
Storage – reduced	10 mM EDTA, 150 mM NaCl, 20 mM HEPES, 2 mM TCEP
Storage – non-reduced	10 mM EDTA, 150 mM NaCl, 20 mM HEPES

Table 2.4

2.1.7.3 Buffers for assessing the locked toroid structure of Prx

Buffer	Composition
Non-reduced	150 mM NaCl, 20 mM HEPES
Reduced	150 mM NaCl, 20 mM HEPES, 2 mM TCEP
EDTA chelation (red/NR)	10 mM EDTA, 150 mM NaCl, 20 mM HEPES, ± 2 mM TCEP
Imidazole chelation (red/NR)	20 mM imidazole, 150 mM NaCl, 20 mM HEPES, ± 2 mM TCEP
EDTA + imidazole chelation (red/NR)	10 mM EDTA, 20 mM imidazole, 150 mM NaCl, 20 mM HEPES, ± 2 mM TCEP

Table 2.5 (NR – non-reduced, red - reduced)

2.1.7.4 Buffers for metal directed oligomerisation of Prx

Buffer	Composition
Association buffer	20 mM imidazole, 150 mM NaCl, 20 mM HEPES, pH 8.0
Coordination buffer	150 mM NaCl, 20 mM HEPES, pH 8.0
Association buffer (Ardini) [4]	20 mM imidazole, 150 mM NaCl, 30 mM MOPS, pH 7.6
Coordination buffer (Ardini) [4]	150 mM NaCl, 30 mM MOPS, pH 7.6
Chelation buffer	50 mM EDTA, 150 mM NaCl, 20 mM HEPES, pH 8.0

Table 2.6

2.1.7.6 pH variation buffers for the oligomerisation of Prx

Buffer	Composition
pH 7.8 – 7.0 red/NR	150 mM NaCl, 50 mM HEPES, ± 2 mM TCEP
pH 9.0	150 mM NaCl, 50 mM Tris base
pH 4.0	150 mM NaCl, 50 mM citrate
High salt A	800 mM NaCl, 20 mM HEPES, pH 8.0
High Salt B	800 mM $(NH_4)_2SO_4$, 20 mM HEPES, pH 8.0
High salt C	300 mM NaCl, 50 mM HEPES, pH 7.4
High salt D	600 mM NaCl, 50 mM HEPES, pH 7.4
High salt E	1.2 M NaCl, 50 mM HEPES, pH 7.4
No salt A	20 mM HEPES, pH 8.0
No salt B	50 mM HEPES, pH 7.4

Table 2.7

2.1.7.5 Peroxidase activity assay buffer

Buffer	Composition
Activity assay buffer	100 mM KH_2PO_4, 0.1 M DTPA, pH 7.5

Table 2.8

2.1.7.5 TEV purification and cleavage buffers

Buffer	Composition
Lysis buffer	20 mM imidazole, 20 mM NaH_2P_i, 500 mM NaCl, 0.3 mM TCEP, pH 7.5
Elution buffer	500 mM imidazole, 20 mM NaH_2PO_4, 500 mM NaCl, 0.3 mM TCEP, pH 7.5
Storage/cleavage buffer	20 mM NaH_2PO_4, 150 mM NaCl, 1 mM TCEP, 1 mM EDTA, pH 7.5

Table 2.9 *NaP_i relates to the ratio of NaH_2PO_4 and Na_2HPO_4 that has been calculated to achieve the correct pH.* [3]

2.2 Molecular biology and microbiology methods

2.2.1 Preparing competent cells

The frozen glycerol stocks of BL21(DE3)*or BL21(pRARE)Rosetta cells were applied to an agar plate containing either no antibiotic (BL21(DE3)*) or chloramphenicol (BL21(pRARE)Rosetta) (Table 2.1) (Chapter Two at 2.1.6.2) using a sterile loop. [1] The plates were incubated overnight (O/N) at 37^{o}C at which point a single colony was picked for a 10 mL LB seed culture. The seed culture (± chloramphenicol) was grown O/N with shaking at 37 ^{o}C. 1 L of LB media (Chapter Two at 2.1.6.1) was inoculated with seed culture (1 %) (± chloramphenicol) and grown at 37 ^{o}C until the optical density at 600 nm (OD_{600}) reached between 0.3 – 0.4. The solution was cooled to 4 ^{o}C on ice and the cells were harvested via centrifugation (3000 x *g*, 15 mins, 4^{o}C). At this point the cells were prepared for either electroporation or chemical transformation. [5]

2.2.1.1 Electrocompetent cells

After centrifugation the supernatant was decanted and the pellet resuspended in 1 L of ice cold, sterile water. The cells were harvested from the suspension via centrifugation (3000 x *g*, 15 mins, 4^{o}C). The supernatant was removed and the pellet resuspended in 500 mL of ice cold, sterile water. The cells were harvested using centrifugation. *These wash steps are essential as they remove any salts from the pellet that may interfere with the electroporation process.* Again, the cells were resuspended in 20 mL ice cold 10 % glycerol and the centrifugation process was repeated. At this point, the supernatant was removed and the cell pellet was resuspended 4 mL ice cold 10 % glycerol. The suspension was separated into aliquots of 100 μL, frozen in liquid nitrogen and stored at – 80^{o}C until use.

2.2.1.2 Chemically competent cells

After centrifugation the pellet was resuspended in 400 mL, ice cold, sterile $MgCl_2$ solution (100 mM). The cells were harvested via centrifugation (2000 x *g*, 15 mins, 4^{o}C). The supernatant was discarded and the pellet resuspended in 200 mL of ice cold, sterile $CaCl_2$ and incubated on ice for 20 minutes. Again, the cells were harvested via centrifugation (2000 x *g*, 15 mins, 4^{o}C) and the supernatant removed. The pellet was resuspended in 50 mL ice cold, sterile chemically competent cell storage solution (85 mM $CaCl_2$, 15 % glycerol w/v). The centrifugation step was repeated again but at a slower velocity (1000 x *g*, 15 mins, 4^{o}C). The supernatant was discarded and the cells resuspended in 2 mL of the chemically competent cell storage solution. The suspension was separated into aliquots of 100 μL, frozen in liquid nitrogen and stored at – 80^{o}C until use.

2.2.2 Competent cell transformation

2.2.2.1 Electrocompetent cells

100 μL of frozen cells were thawed gently on ice (4^{o}C). 1 μL of plasmid solution was added to the cells and mixed gently by tapping the Eppendorf and the mixture was incubated on ice (4^{o}C) for 30 minutes. The mixture was transferred to a pre-chilled, sterile 1 mm electroporation cuvette and placed into the elctroporator (BioRad Gene PulserTM). It was essential that the cuvette was cleaned thoroughly prior to transformation to prevent unwanted

DNA contamination. The sample was subjected to a 2.5 kV pulse and 900 μL of sterile LB (no antibiotic) was added immediately. The mixture was incubated at 37^{o}C for one hour. Agar plates, containing the appropriate antibiotics (Table 2.1 and 2.2) were inoculated with the transformed cells under sterile conditions, [1] and incubated O/N at 37^{o}C.

2.2.2.2 Chemically competent cells

100 μL of frozen cells were thawed on ice (4^{o}C). 1 μL of plasmid solution was added to the cells and mixed gently by tapping the Eppendorf. The mixture was incubated on ice for 30 minutes before heat shock at 42 ^{o}C for 1 minute. The cells were immediately put back on ice for a further 2 minutes. At this point, 900 μL of sterile LB media was added to the aliquot and the sample was mixed by pipetting and the mixture was incubated at 37 ^{o}C for one hour. Agar plates, containing the appropriate antibiotics (Table 2.1 and 2.2) were inoculated with the transformed cells under sterile conditions, [1] and incubated O/N at 37 ^{o}C.

2.2.3 Bacterial culture storage

2.2.3.1 Long term storage

Glycerol stocks were made for long term culture storage. 10 mL of LB (Chapter Two at 2.1.6.1), + the appropriate antibiotic (Table 2.1), was inoculated with a single colony of transformed cells under sterile conditions. [1] This was grown O/N at 37^{o}C with shaking (180 rpm). 2 mL of seed culture was diluted 50:50 with sterilised glycerol solution (50 %) for a final glycerol concentration of 25 %. The mix was split into aliquots (1 mL), flash frozen with liquid nitrogen and stored at -80^{o}C indefinitely.

2.2.3.1Short term storage

The cells were grown on agar plate prepared using the protocol described in 2.1.6.2. The plates were incubated O/N at 37^{o}C until mature colonies formed (~16 hours). The sealed plates were stored for up to 2 weeks at 4^{o}C.

2.2.4 DNA manipulations

2.2.4.1 Plasmid extraction

Reagents	Composition
glucose-Tris-EDTA (GTE)	50 mM glucose, 25 mM Tris-HCl, 10 mM EDTA, pH 8.0
NaOH/SDS solution	0.2 M NaOH, 1% SDS
potassium acetate solution	5 M CH_3O_2K
Isopropyl alcohol (IPA)	ACS reagent ≥ 99.5%
Ethanol (EtOH)	ACS reagent grade
Tris-HCl-EDTA (TE)	10 mM tris-HCl, 1 mM EDTA, pH 7.5

Table 2.10

Plasmid extraction was performed to ensure correct transformation and to replenish any low plasmids stocks. The methods used were based those developed by Birnboim *et al.* [6] 10 mL LB (Chapter Two at 2.1.6.1) seed culture with the corresponding antibiotics (Table 2.1) was grown O/N at 37^{o}C, with shaking. 1 mL of the seed culture was centrifuged (13, 000 x *g*, 4^{o}C) and the supernatant discarded. 100 µL of GTE buffer was added to the cell pellet. The solution was mixed gently to resuspend the pellet. At this point 200 µL of NaOH/SDS solution was added to the mix and the Eppendorf was inverted 6 – 8 times. As soon as this was completed 5 M potassium acetate solution was added to neutralise the NaOH and precipitate the genomic DNA. To separate out the genomic DNA, the suspension was centrifuged for 2 minutes (13,000 x *g*, 4 ^{o}C, Eppendorf). The supernatant was transferred into a new Eppendorf tube, taking care not to transfer any white flakes. The plasmid DNA was precipitated with the addition of 0.5 mL of IPA. The mix was incubated on ice for 10 minutes and then centrifuged for 2 minutes (13,000 x *g*, 4 ^{o}C). The supernatant was discarded and the sample was left for a few minutes to allow the remaining IPA to evaporate. The pellet was resuspended in 0.5 mL EtOH and the sample was again centrifuged for 2 minutes. The supernatant was discarded and the sample left to air dry, on ice, for 15 minutes. At this point, the pure plasmid was dissolved in 20 µL TE buffer and stored at -20^{o}C until further use.

2.2.4.2 Restriction digest

Restriction digest was used to confirm the plasmids base pair count. The DNA (1 μg) was digested with the appropriate restriction enzymes and corresponding restriction buffers (Chapter Two at 2.1.5.2). The products were analysed on agarose electrophoresis (Chapter Two at 2.2.4.4)

2.2.4.3 DNA concentration

The concentration of DNA samples was analysed using NanoDrop 3300 spectrometer (Thermo Fisher Scientific Inc) using the automatic function:

$$\text{Concentration} = (A_{260} / \varepsilon_{260}) * l$$

A_{260} = absorption at 260 nm (AU)

ε_{260} = wavelength dependent extinction coefficient at 260 nm (ng/cm/μL)

l = path length

The software calculated the concentration from the average extinction coefficient for single stranded DNA which is around 50 ng/cm/μL.

2.2.4.4 Agarose gel electrophoresis

50 X Tris-Acetate EDTA (TAE) buffer ~ pH 8.5, 1L	
Tris-base	242 g
Glacial acetic acid	57.1 mL
EDTA (0.5 M, pH 8.0)	100 mL

Table 2.11

Gel electrophoresis was carried out to check DNA purity, identify DNA contamination from purified Lsr2 and to ensure the presence of the correct primers. Agarose gels were prepared by adding 1 % agarose (w/v) to 1 X TAE buffer and dissolving by gentle heating, ensuring

the solution did not boil. Upon cooling to < 50^{o}C, 5 μL of SYBR® Safe dye (Thermo fisher Scientific, Waltham US) was added and the gel was set in a gel caster, which included a 10 well comb. Once the gel had set the cassette was placed into an electrophoresis chamber (Mini-sub Cell GT Cell, Bio-Rad, Hercules, US) and covered with 1 X TAE buffer until the liquid just filled the wells. DNA samples were mixed with a 1/6 volume of 6 x DNA loading dye (Life Technologies, Carlsbad US) to a total volume of 12 μL. 10 μL of this mix was loaded into the wells, taking care not to load any air bubbles. 5 μL of DNA ladder (various lengths, Life Technologies, Carlsbad US) was loaded into the remaining well and the sample was run at 120 V for 1 hour. The loading buffer contained a fluorescent dye and the bands corresponding to DNA base pairs were visualised under UV light.

2.2.4.5 DNAse digest

10 X DNase buffer pH 7.6	
Tris-HCl	100 mM
$MgCl_2$	25 mM
$CaCl_2$	5 mM

Table 2.12

DNAse was used to digest the DNA that co-purified with Lsr2. [7] 3 μL of 10 X DNAse buffer was added to 25 μL protein sample (1 mg/mL) and 2 μl of RQ1 DNase (1 U/μl, Promega) was mixed in to a total volume of 30 μL. The digest was left for 90 minutes at room temperature and the products were analysed on agarose gel electrophoresis (Chapter Two at 2.2.4.4)

2.3 Recombinant protein production

2.3.1 Protein expression

Agar plates (Chapter Two at 2.1.6.2), supplemented with the appropriate antibiotic (Table 2.1 and 2.2) were inoculated with ~ 2 μg of glycerol stock culture under sterile conditions. [1] This was spread evenly across the plate using a sterile loop. The plates were incubated O/N at 37^{o}C until established colonies appeared. A single colony was used to inoculate a 10 mL LB seed culture, supplemented with the correct antibiotics (Table 2.1 and 2.2). This was grown O/N, with shaking (180 rpm) at 37^{o}C

2.3.1.1 Isopropyl β-D-1-thiogalactopyranoside (IPTG) induction

Lysogeny broth (LB) was prepared using the methods in 2.1.6.1. Either 1L (Lsr2, Nterm-Lsr2, Ent-Nterm) or 400 mL (HsPrx3 and its derivatives) was inoculated with 1 % seed culture + the appropriate antibiotics (Table 2.1 and 2.2). The broth was incubated at 37^{o}C, with shaking (180 rpm). The optical density at 600 nm (OD_{600}) was monitored on a Specta Max MS multi-mode microplate reader at set time points until it reached between 0.4 – 0.6. At this point, the culture was cooled to room temperature and IPTG was added to a final concentration of 600 μM. The culture was grown between 16 – 20 hours, with shaking, at 26^{o}C

2.3.1.2 Auto induction

Media was prepared following the methods in 2.1.6.3. 1 L of media was inoculated with 1 mL of seed culture + the corresponding antibiotics (Table 2.1 and 2.2). The broth was grown, with shaking (180 rpm), for 4 hours at 37^{o}C. At this point the temperature was reduced to 30 ^{o}C and the culture was grown for a further 16 hours with shaking (180 rpm).

2.3.1.3 Cell harvest by centrifugation

The cell culture was harvested via centrifugation (8000 x *g*, 20 minutes, 4^{o}C). The pellet was weighed and lysis buffer (Table 2.3 and 2.4) was added at a ratio of 5 x w/v. Once resuspended, the mixture was either lysed immediately or stored at – 80^{o}C for up to 2 weeks.

2.3.2 Protein extraction and purification

2.3.2.1 Small Scale Lysis

Small scale lysis was completed to confirm successful protein expression. 1 mL of the cell culture was harvested by centrifugation (8000 x *g*, 5 mins, RT). The pellet was resuspended in 200 µL of lysis buffer (Table 2.3 and 2.4) by gentle pipetting. 20 µL of 10 x lysozyme (10 mg/mL) was added to the crude lysate along with Triton X-100 ASR (detergent) to a final concentration of 0.1%. The mixture was incubated in a water bath, with shaking, for 30 minutes at 37^{o}C. To aid cell lysis the lysate was subjected to a freeze thaw cycle (freeze in liquid nitrogen and thaw at 37^{o}C for 15 min). Once fully thawed, 20 µL of 10 x DNase (1 mg/mL) and 20 µL of 10 x $MgSO_4$ (1 M) were both added and the sample was shaken at RT for 15 minutes. The insoluble and soluble fractions were separate by further centrifugation (10 000 x *g*, 5 mins, 4^{o}C). The concentration of the soluble crude was analysed using nanodrop (Chapter Two at 2.4.1.2) and the presence of the target protein was confirmed by SDS-PAGE (Chapter Two at 2.4.2).

2.3.2.2 Large Scale Lysis

The resuspended pellet was cooled on ice to 4^{o}C. To this mix, lysozyme was added to aid the lysis process (1 mg/mL). Cells were lysed on ice using sonication (0.5 amplitude, 40 %, 5 minutes x 4) with one minute resting between each step. This was to allow the sample to cool to 4^{o}C and prevent over heating which could denature the protein.

2.3.2.3 Soluble protein extraction

To separate the soluble proteins from the insoluble proteins and cell debris the sample was hard spun at 4^{o}C (18000 x *g*, 30 minutes). At this point the soluble crude fraction was separated from the insoluble pellet, filtered through a 0.22 µm syringe filter, and kept on ice for further purification.

2.3.2.4 Protein extraction by immobilised metal ion chromatography (IMAC)

All constructs had an N-terminal or C-terminal 6-histdine tag. This allowed for IMAC purification. The target protein binds to immobilised Ni^{2+} within the column via the histidine imidazole side chain (~10 mg/mL, protein/column volume (CV)), and is released upon an increasing external imidazole concentration. The filtered soluble crude fraction was loaded onto an IMAC column (HisTrap FF Crude 5 mL column, GE Healthcare) equilibrated with lysis buffer (Table 2.3 and 2.4), at a flow rate of 2 mL/min. The column was then washed with 10 x CV or until all of the non-bound or weakly bound protein had eluted (absorbance at 280 nm returns to baseline level). Any flow through (unbound protein) was collected to ensure the target protein had coordinated to the column. The target protein was eluted by applying a linear or stepwise gradient of an elution buffer (Table 2.3 and 2.4). Eluted fractions were collected in a 96 well plate and the absorbance signal at 280 nm was used to identify the location of the target protein. Fractions containing the target protein were pooled and concentrated to 5 mL using a 3 or 10 kDa molecular weight cut off concentrator (Centrifugal Concentrators Vivaspin 20 3.000 or 10.000 Cut Off), depending on the calculate MW of the protein.

2.3.2.5 Buffer exchange by desalting column

Isolated target protein from IMAC was passed through a desalting column (HiTrap Desalting 5 mL column, GE Healthcre), equilibrated with protein storage buffer, (Table. 2.3 and 2.4) to remove excess imidazole. Eluted fractions were collected in a 96 well plate. The exchange of buffer was confirmed by measuring the conductivity and absorbance at 280 nm. A peak in the conductivity relates to the eluted salt. In successful buffer exchange there was no cross over between the conductivity peak and 280 nm absorbance peak.

2.3.2.6 Size exclusion chromatography (SEC)

SEC was used during the purification as a final finishing step and in some cases, as a buffer exchange step. (GE Healthcare HiLoad 10/16 Superdex 200 prep grade , MW 10 – 600 kDa) Due to the large volume of the column, up to 5 mL of sample could be loaded (< 3 % CV) [8] and the Superdex matrix allowed for high resolution separation which made this column ideal for purification purposes. [9] Blue dextran was eluted through the column to assess the void

volume (V_0), calculated to be 46.1 mL. Before sample analysis and purification, the column was calibrated using commercially available standards of a known MW (Table 2.13). The partition (K_{av}) was calculated from the standard elution volumes (equation 2.1).

$$K_{av} = (V_e - V_0)/(V_c - V_0)$$

V_e = elution volume

V_0 = void volume (46.1 mL)

V_c = column volume (120 mL)

Equation 2.1

Log(MW) was plotted against K_{av} to give a calibration curve from which an unknown MW could be calculated (equation 2.2).

$$MW = 10^{((0.34 \times Kav) + (1.20)}$$

Equation 2.2

Protein	MW (Da)	V_e	Log(MW)	Kav
Thyroglobulin	670,000	45.622	5.826075	-0.00647
γ-globulin	158,000	64.003	5.198657	0.24226
Ovalbumin	44,000	81.326	4.643453	0.476671
Myoglobin	17,000	92.699	4.230449	0.630568
Vitamin B12	1,350	112.138	3.130334	0.893613

Table 2.13: Elution volumes of calibration standards of a known molecular weight. Log(MW) is plotted against K_{av} to give the HiLoad 10/16 Superdex 200 prep grade calibration curve.

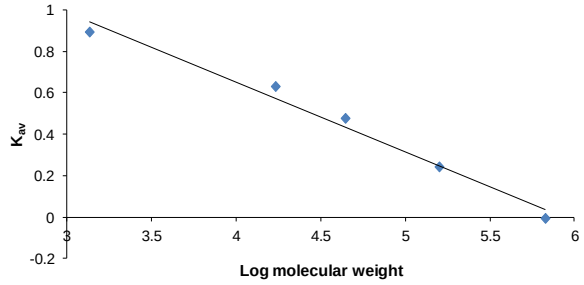


Figure 2.1: Calibration curve for the HiLoad 10/16 Superdex 200 prep grade column.

The column was equilibrated with storage buffer (Table 2.3 and 2.4) and the sample was loaded and run at 1 ml/min. The absorbance (280nm) was measured to allocate protein elution volumes. [8] The MW of the eluted fractions was calculated using the calibration curve. [10] All fractions were analysed on SDS-PAGE (Chapter Two at 2.4.3) to confirm the presence of target protein/the absence of impurities. Fractions containing isolated target protein were stored at 4 ^{o}C (short term, 1 week) or flash frozen in liquid nitrogen (5 – 10 %) and stored at -80 ^{o}C until use.

2.3.3 Histidine tag removal by Tobacco etch virus (TEV) digest

2.3.3.1 TEV expression and purification

TEV was produced in-house using the plasmid design (pRK793) derived from Blommel and Fox 2007. [11] An affinity tag, poly-6histidine, followed by a maltose binding protein (MBP) was added to the C-terminus of the target protein by expression vector. This allowed ease of purification and increased protein solubility. [11] There was a TEV recognition site between the MBP and histidine tag which allowed for self-cleavage of the MBP during the purification process. *E. coli* (BL21 RILP) was transformed with the plasmid and the cells were grown on agar plates supplemented with kanamycin and chloramphenicol. The protein was expressed in auto induction media (Chapter Two at 2.1.6.3) using the protocols from section 2.3.1. [2] The cells harvested using centrifugation, resuspended in lysis buffer (Table 2.9) and lysed on ice via sonication (Chapter Two at 2.3.2.2). The cell debris and insoluble protein was removed by centrifugation (Chapter Two at 2.3.2.3). Before any further purification, the sample was left to incubate at 4 ^{o}C for at least 3 hours to allow for the self-cleavage of MBP. [11] Target protein was isolated from the soluble crude fraction using IMAC (Chapter Two at 2.3.2.2) and eluted upon increasing elution buffer concentration (Table 2.9). SEC was used a final purification step and to exchange the protein into storage/cleavage buffer (Table 2.9). Isolation of the target protein was confirmed by SDS-PAGE (Chapter Two at 2.4.3).

2.3.3.2 TEV fusion tag cleavage

TEV cleavage was used to remove the histidine tag from *Hs*Prx3-6his. The protein had a TEV recognition site which allowed for the removal of the non-native residues at the N-terminus. A small scale digest using a range of TEV:*Hs*Prx3-6his ratios was completed to determine the optimal conditions for tag removal

*Hs*Prx3-6his was transferred into the cleavage buffer (Table 2.9) using a desalting column. The protein was then mixed with TEV at ratios 1:1, 1:10, 1:20, 1:50 and 1:100. The samples were incubated O/N at 4 ^{o}C. The extent of digest was analysed using SDS-PAGE (Chapter Two at 2.4.3). Successful digest at 1:20 TEV:*Hs*Prx3 was confimed by a single band at 20 kDa (cleaved protein MW). 2 mL of *Hs*Prx3-6his (8 mg/mL) was transferred into cleavage buffer using a desalting column. The eluted protein was diluted to 2 mg/mL and split into aliquots of ~ 2mL. TEV was added to a final ratio of 1:20 and the samples were incubated O/N at 4 ^{o}C. At this point, the aliquots were combined and cleaved protein was isolated from the tag and TEV using IMAC that had been charged with Co^{2+}. The pass through was collected and further purified using SEC. This was also the desalting step whereby the cleaved protein was transferred into storage buffer (Table 2.4). The purity of the cleaved protein was confirmed using SDS-PAGE.

2.4 Protein characterisation

2.4.1 Determining protein concentration

2.4.1.1 Bradford assay

Reagents for Bradford assay	
Dye	Bio-rad Protein Assay Dye Reagent concentrate (Bio-Rad)
Protein standard	Bovine Serum Albumin (BSA) (Gibco)

Table 2.13

Bradford assay was used to determine protein concentration following procedures developed by Marion M. Bradford. [12] BSA standards were prepared in dH_2O with a concentration range from 0.50 – 0.063 mg/mL. The protein samples were analysed at 100%, 50 % and 10 % dilution. The upper limit of the wavelength with the Bradford reagent is 0.5 mg/mL, anything above this will not be accurate and therefore the dilutions are needed to ensure accurate measurement. Samples were prepared in a 96 well, clear bottom plate (Jet Biofil). 10 µL of the protein dilution was added to each well and then mixed with 200 µL of 1 x Bradford reagent (Biorad, Hercules, CA). All standards and protein dilutions were prepared in triplicate. The plates were left to stand for at least 5 minutes but no longer than hour, at room temperature to allow the dye to develop. The concentration was analysed using a Specta Max MS plate reader (λ = 595 nm). Protein concentrations was calculated from the BSA calibration curve.

2.4.1.2 Nanodrop

For samples with a low volume, nanodrop was used to gain an accurate concentration. This was compared to the bulk sample to ensure method consistency. The nanopdrop ND 100 spectrophotometer was equilibrated with 2 µL of filtered dH_2O and the equipment was blanked using the appropriate buffer. The protein was loaded on the reader and measured the 260/280 nm using the protein extinction coefficient calculated from ExPASy. [13]

2.4.2 Differential scanning fluorometry

DSF reagent	
50 X dye	2 µL 5000 X SYPRO orange (Invitrogen) 198 µL dH_2O

Table 2.14

The 50 X dye was prepared and centrifuged (12, 000 rpm, 1 min, Eppendorf) directly before use. The protein sample (40 µL, ~ 1 mg/mL), dH_2O (50 µL) and 50 X dye (10 µL) were mixed to a final volume of 100 µL. The mixture (25 µL) was added to a 96-well thin wall, PCR plate (Bio-Rad) in triplicate. A buffer blank was also prepared to generate the thermal

shift base line. The plates were sealed and heated on a Biorad, CFX connect Real Time PCR Detection System (Bio-Rad) from 20 to 95 °C in increments of 0.5 °C, with 30 s dwell time. The melting temperature was calculated by plotting the negative derivative of the relative fluorescent units (-*d*(RFU) *d*t) against temperature (^{o}C).

2.4.3 Sodium Dodecyl Sulfate Polyacrylamide gel Electrophoresis (SDS-PAGE)

SDS-PAGE reagents	
4 x Loading buffer	Novex® Sharp Pre-stained Protein Standard (Life Technologies)
10 x reducing agent	500 mM DTT
Molecular weight ladder (MWL)	Novex® Sharp Pre-stained Protein Standard (Life Technologies)
Gel	NuPAGE™ Novex™ 4-12% Bis-Tris Protein Gels, 1.0 mm, 10-well (Life Techonolies, NP0321)
Electroporesis tank buffer	NuPAGE® MES SDS Running Buffer (20X) (Life Technologies)

Table 2.15: List of reagents, buffers and equipment needed for SDS-PAGE

SDS-PAGE analysis was used routinely to determine protein molecular weight and to probe the stability of oligomers. [14] Samples were prepared with 5 µL 4x loading buffer, 2 µL 10 x DTT, protein solution to a final concentration of around 250 µg/mL and dH_2O to a final volume of 20 µL. To ensure protein denaturation, the aliquots were placed in nearly boiling water (95 ^{o}C) for 5 minutes and then centrifuged before analysis. The 20 µL samples were loaded onto the gel and run at 200 mV for 35 minutes. The gel was washed in warm dH_2O 3 times to allow any SDS to defuse from the gel and stained with simply blue stain (60 – 80 mg

G250) for at least 2 hours or until clear bands began to appear. Destaining was achieved through O/N rinsing in dH_2O.

2.4.4 Matrix assister laser ionisation (MALDI)-time of flight (TOF) mass spectrometry

Solutions for MALDI-TOF MS

Matrix diluent (1 mL)	500 µL acetonitrile 500 µL uH_2O 1 µL triflouroacetic acid
Sample/standard diluent (1mL)	300 µL acetonitrile 700 µL uH_2O 1 µL triflouroacetic acid
Matrix HMW	Sinapinic acid
Matrix LMW	α-cyano-4-hydroxycinnamic acid (CHCA)
Cal 1 LMW (100 µL) **(Applied Biosystems)**	6.5 µg Angiotensin I 10.5 µg Adrenocorticotropic Hormone (1–17 clip) (2094Da) 9.3 µg Adrenocorticotropic Hormone (18–39 clip) (2464 Da) 27.5 µg Adrenocorticotropic Hormone (7–38 clip) (3659 Da) 50.2 µg Insulin, bovine (5733 Da)
Cal 2 HMW (100 µL) **(Applied Biosystems)**	7.1 µg Insulin, bovine (5733 Da) 80.0 µg Thioredoxin, *E. Coli* (11674 Da) 170.0 µg Apomyoglobin, horse (16,952 Da)

Table 2.16

The MW of Lsr2, Nterm-Lsr2 and Ent-Nterm pre and post-proteolysis was ascertained using MALDI-TOF mass spectroscopy. For the proteins pre-proteolysis the HMW standard and matrix were used and for the proteins post-proteolysis the LMW matrix and standards were used. The matrix (HMW or LMW) was mixed in 1 mL of matrix diluent directly before use.

Undissolved matrix crystals were removed by centrifugation (12, 000 rpm, 5 mins, Eppendorf). The sample matrix and protein sample/calibration standard were mixed at a 5:1 ratio. 1 µL of this mix was spotted in duplicate onto the MALDI-TOF sample plate (192 well, Applied Biosystems) and air dried before analysis.

Spectra was obtained on an AB SCIEX 5800 MALDI TOF/TOF at Victoria University, Wellington in the positive ion linear mode. The parameters were as follows:

Detector voltage – 1.95 kV
Pulse rate – 400 Hz
Time delay – 200 - 300 ns
Mass range – 2000 – 30,000 Da

2.4.5 Liquid Chromatography Mass Spectroscopy (LCMS)

LCMS was used to determine the protein MW for constructs > 20 kDa. The protein was exchanged into a LCMS buffer using the desalting column. 0.5 mL aliquots of ~2 mg/mL protein sample were transferred into a 1 mL glass vial in preparation for analysis. The mass was obtained on a maXis 3G ultra high resolution time of flight tandem mass spectrometer (Brucker, Billerca, US). All data collection and analysis was completed by Dr Amelia Albrett at the University of Canterbury.

2.4.6 Western blot immunodetection

Western blot immunodetection was used to identify bands on SDS-PAGE that related to SDS stable target protein oligomers. The 1^{o} antibody (anti-mouse) binds to the histidine tag sequence [15] and a 2^{o} antibody is applied which binds to the 1^{o} and the chromogenic dye which allows for tagged protein detection. The samples to be analysed were run on SDS-PAGE (Chapter Two at 2.4.3). All solutions were part of the anti-mouse iBlot chromogenic western detection kit (Invitrogen, Carlbad, US) The diluent mix (10.5 mL of diluents + 4.5 mL of additive) was prepared and split into three aliquots 5 mL prior to protein transfer.

2.4.6.1 Protein transfer

The protein from the gel was transferred to a PVDF membrane using an iBlot transfer system (Invitrogen, Carlbad, US) The PVDF membrane was placed on a copper anode (bottom stack). The SDS-PAGE gel was removed from the cassette and placed on top of the PVDF membrane followed by a sheet of filter paper, soaked in dH_2O, and a copper cathode (top stack). Air bubbles were gentle removed using a roller and a sponge was placed on top of the cathode. the lid was closed to complete the circuit. Program 2 (23 V, 6 mins) was selected to apply an electric field which transferred the charged protein to the membrane.

2.4.6.2 Blocking

The PVDF membrane was washed with dH_2O and then placed back on top of the copper anode. The equipment was set up as described above with a blank matrix (soaked in 5 mL of diluent mix) in place of the filter paper. Program 9 was selected which ran in three stages:

1. 20 V, 2 mins
2. 5 V, 3 mins
3. 5 V, 3mins

While the first blocking step was running a new matrix was soaked in 5 mL of 1^o antibody mix (5 mL diluent mix + 5 µL 1^o antibody). The 1^o antibody matrix replaced the blank matrix and the second stage for program 9 was run. A new matrix was soaked in 5 mL 2^o antibody mix (5 mL diluent mix + 5 µL 2^o antibody) and the procedure was repeated for the final step.

2.4.6.3 Washing

The PDVF membrane was briefly rinsed with 20 mL of 1 x wash solution (6 mL stock washing solution + 90 mL dH_2O). The solution was discarded and another 20 mL of 1 x wash solution was added and the sample was incubated at RT (5 mins with gentle shaking). This step was repeated two more times, discarding the wash solution at each step. The membrane was rinsed twice with dH_2O.

2.4.6.4 Chromogenic detection

The membrane was soaked in 5 mL of the chromogenic substrate solution. This was incubated until a purple band for the histidine tagged positive control began to appear (no

longer than 1 hour). The reaction was stopped by rinsing with dH_2O for 2 minutes x 2. The membrane was air dried and the image recorded.

2.5 Biophysical techniques

2.5.1 Analytical SEC

Analytical SEC was completed on a Superdex 200 10/300GL, 24 mL (GE Healthcare), MW range 10 – 600 kDa. The Superdex matrix, dextran beads covalent linked to cross-linked agarose, allowed for high resolution analysis and the smaller volumes analysed using this column as appose to the 120 mL column reduced the peak broadening effect that can be a hindrance when using this method of MW analysis. [8] Blue dextran was eluted through the column to assess the void volume (V_0), calculated to be 8.34 mL. Before sample analysis and purification, the column was calibrated using commercially available standards of a known MW (Table 2.17). The partition (K_{av}) was calculated from the standard elution volumes (equation 1, 2.3.2.6, V_c = 24 mL). Log(MW) was plotted against K_{av} to give a calibration curve from which an unknown MW could be calculated (equation 2.3).

$$MW = 10^{((0.30xKav) + (1.84)}$$

Equation 2.3

Protein	MW (Da)	V_e	K_{av}	log MW
Thyroglobin	669000	9.181	0.055874	5.825426
Ferritin	440000	10.935	0.167622	5.643453
Aldolase	158000	13.17	0.310015	5.198657
Conalbumin	75000	14.525	0.396343	4.875061
Ovalbumin	44000	15.547	0.461455	4.643453
Carbonic Anhydrase	29000	16.84	0.543833	4.462398
Cytochrome C	12400	18.002	0.617864	4.093422

Table 2.17: Elution volumes of calibration standards of a known molecular weight. Log(MW) is plotted against K_{av} to give the Superdex 200 10/300GL columne calibration curve.

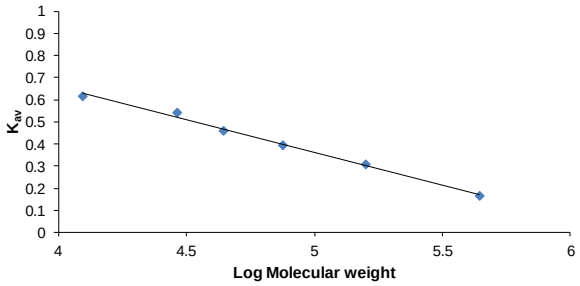


Figure 2.2: Calibration curve for Superdex 200 10/300GL column.

Protein samples were prepared in the appropriate buffer (Chapter Two at 2.1.7) at varying concentrations (0.2 – 10 mg/mL) depending on the experiment. The column was equilibrated in the sample buffer and the protein solution (0.5 mL) was loaded onto the column and run at a flow rate of 0.5 mL/min. The absorbance (280nm) was measured to allocate protein elution volumes. [8] In some cases the eluted fractions were collected for further analysis e.g. equilibrium experiments or SDS-PAGE analysis to check sample purity. The MW was calculated from the column calibration curve. This particular method of analysis allowed for the calculation of protein oligomer MW [10] and therefore was used as a primary method to assess the degree of assembly in the protein native state and under non-native self-associating conditions.

2.5.2 Circular Dichroism (CD)

Circular dichroism was employed to analyse the protein secondary structure. The protein was transferred into 10 mM sodium phosphate buffer to remove any salt (NaCl) and reducing agents (TCEP) that strongly absorb within the range of the wavelength scan. [16] The buffer exchange was completed using a desalting HiTrap 5 mL column (Chapter Two at 2.3.2.7). A quartz cuvette (1 mL) with a 1 cm path length was used for all experiments. The measurements were made using a ChiraScan CD spectrometer (Applied Photophysics Ltd, United Kingdom). The instrument was initialised with dH_2O and the buffer was used to generate a blank spectra. The CD in response to absorption was measure over a range of wavelengths (λ 280 – 190 nm). Correct secondary structure was confirmed by matching the secondary elements on the spectra [16] with the crystal structure of protein. [16] Any

variations in the secondary structure due to changes in the protein sequence or experimental conditions were measured in this way.

2.5.3 Analytical ultra centrifugation (AUC)

Protein solutions were prepared at concentrations ranging from 0.15 – 0.6 mg/mL to a volume of 380 µL. Each sample was analysed with its own blank (400 µL) containing the corresponding buffer only. The larger blank volume was needed to accurately place the meniscus of the sample cell. [17] The cells used were either quartz or sapphire with a cell path length of 1.2 cm. A wavelength scan was run prior to the experiment to determine to wavelength (λ = 270 – 280) that provided the optimal optical density ($OD_{270-290}$ between 0.2 – 1.0 [17]) for all the protein samples that were being analysed at any one point. During this study the centrifugal velocity was initially set at 38 000 rpm as this had shown good data around the MW of a peroxiredoxin dimer and dodecamer. Samples of a larger MW were studied at 20 000 rpm to allow for sedimentation over a greater number of scans. All runs were completed at 20 ^{o}C. The Svendberg sedimentation coefficient (S) was calculated using the c(s) model in SEDFIT. [18] The software was also able to approximate the particle mass using the equation below (equation 2.4)

$$M = \frac{s \cdot f \cdot Na}{1 - \rho \cdot \bar{\upsilon}}$$

M = mass
s = sedimentation coefficient
f = frictional coefficient
N_a = Avogadro's number
ρ = solvent density
$\bar{\upsilon}$ = partial specific volume

Equation 2.4

2.5.4 Transmission electron microscopy (TEM)

The structure of the proteins in their native and oligomerised state were analysed using TEM. The heterogenous nature of many of the oligomer systems would make other structure determining techniques, (X-ray crystallography, SAXS) difficult. The protein samples were

diluted to 100 or 50 μg/mL in dH_2O directly before grid preparation. This was to reduced over crowding from excess protein and to remove salts, both of which would diminish the clarity of the image collection. Carbon coated formvar 200-mesh copper grids were deposited for 1 minute on;

1) 20 μL protein sample
2) 20 μL (x 3, 20 seconds each)
3) 1 % uranyl acetate negative stain

Excess dye was removed by holding fibre less filter paper to the edge of the grid and allowing the liquid to be soaked up, taking care to not touch the centre of the grid. The grid was then air-dried for a least an hour before TEM analysis.

Image collection was completed by Jackie Healey, University of Canterbury, on a FEI Morgagni 268D transmission electron microscope operating at 80 kV, with magnifications up to 180k. Images were captured using a SIS/Olympus Megapixel III digital camera mounted above the phosphor screen.

2.5.5 Dynamic light scattering (DLS)

DLS was used a quick method to analyse the particle size of protein assemblies. It was also used to map the progression of oligomerisation in real time. DLS was measured on a Zetasizer Nano ZS (Malvern UK) which uses a He-Ne laser as the light source ($\lambda = 673$) with the 173 ^{o} back scattering detected by an avalanche photodiode. The proteins were prepared in the appropriate buffer (Chapter Two at 2.1.7) at varying concentrations (0.1 – 1 mg/mL) and the cuvette l (standard disposable cuvette, PMMA, Sigma Aldrige) was placed into the thermostated sample chamber, set at 20 ^{o}C. The sample was equilibrated to the temperature and the D_H was measure with 3 – 6 scans over 4 minutes. The data was analysed using DTS software which plots the correlation function vs time from which the average particle size (Z-ave) and the polydispersity index (PdI) can be calculated. [19] The results were reported as Z-ave (d.nm).

2.6 Manipulations of Lsr2 and its derivatives

2.6.1 Lsr2 trypsin digest assay

Trypsin (≥7,500 units/mg solid) was used to cleave at the N-terminal lysine (Lys^4). 1 µL of the enzyme solution (1 mg/mL) was added to 50 µL protein sample (1 mg/mL) resulting in a final trypsin to protein solution volume ratio of 1:500. The digest was stopped at varying time points (5, 10, 15 and 30 mins) by heat denaturation (5 mins, 96 ^oC). The MW of the protein was analysed using SDS-PAGE (Chapter Two at 2.4.3). The samples were prepared by adding 5 µL of loading dye and 2 µL of DTT solution (0.5 M) to 10 µL of the protein sample and the aliquot was made up to 20 µL with dH_2O. Trypsin was also run as a control. The DNA agarose gel samples were prepared and run using the methods in 2.2.4.4, with trypsin also being analysed as a negative control.

2.6.2 Lsr2 in high salt

Lsr2 (1 mg/mL) was transferred into high salt buffer (Table 2.3) using a desalting column (Chapter Two at 2.3.2.5) and left O/N to allow for DNA dissociation. The protein MW was analysed using analytical SEC (Chapter Two at 2.5.2).

2.6.3 Nterm-Lsr2 Trypsin digest

Trypsin (≥7,500 units/mg solid) was used to cleave at the N-terminal lysine (Lys^4) which facilitated the formation of large oligomeric Nterm-Lsr2 structures. [7] 10 µL of the enzyme solution (1 mg/mL) was added to 0.5 mL protein sample of varying concentration resulting in a final trypsin to protein solution volume ratio of 1:500. The digest was stopped after 10 minutes at room temperature by the addition of trypsin inhibitor at 10x the concentration of trypsin. The mixture was left to rest for a further 20 minutes to allow for oligomer formation and injected onto a 24 mL SEC column equilibrated with storage buffer. A small aliquot was retained for TEM analysis.

2.6.4 Ent-Nterm enteropeptidase proteolysis

Porcine-enteropeptidase **(**Applied Biological Materials (ABM) Inc, Canada) was used to cleave at the enteropeptidase linker to leave Lys^4 as the N-terminal amino acid on Ent-Nterm. 30 units of enzyme (30 µL enteropeptidase solution at 1 U/ µL) were added to a 0.5 mL protein sample (1 mg/mL, 50 mM sodium phosphate, 200 mM NaCl, pH 7.4) and stored for 16 h at room temperature. The mixture was then injected onto a 24 mL SEC column equilibrated with storage buffer with a small aliquot kept aside for TEM analysis.

2.6.5 Equilibrium assay

Equilibrium assays were performed on SEC by selecting a fraction from the centre of each oligomer peak and re-injecting onto the column. The concentration for each fraction was measured by the Bradford assay [12] prior to re-injection.

2.6.8 pH sensitivity assay

For variations in pH during oligomerisation, the protein was digested as described above at pH 7.4 followed by the addition of 0.5 mL of a buffer with a 4-fold greater buffer capacity than that of pH 7.4. The pH values tested were 6, 8 and 9 (Table 2.3).

After adjusting and checking the pH, the 1 mL sample was injected onto a 24 mL SEC column equilibrated with the same buffer used to change the oligomerisation pH. As a control a 1 mL sample of Nterm-Lsr2 in pH 7.4 buffer was run under pH 7.4 conditions prior to the pH experiment to show that increasing the sample injection volume did not affect the positioning of peaks on the spectra. Once the sample had been run on SEC the oligomer peaks were collected exchanged into pH 7.4 buffer using a 3 kDa molecular weight cut-off concentrator (Centrifugal Concentrators Vivaspin 20 3.000 Cut Off). Around 3 wash cycles removed 99% of initial salt concentration. The pH of the resulting sample was measured to ensure adequate buffer exchange and the sample was reanalysed on SEC.

2.7 Peroxiredoxin manipulations

2.7.1 Reduced and non reduced

The protein was transferred into the appropriate buffer (Table 2.5) using a desalting column. The concentration was checked by Bradford assay .[12] and a small aliquot was retained for SDS-PAGE analysis. 0.5 mL was loaded onto the SEC analytical column and the MW was calculated from the calibration curve (Chapter Two at 2.5.1). The samples for SDS-PAGE were denatured using slightly altered conditions to the standard protocol (3.4.3) (95 ^{o}C, 1 min) and the oligomerisation was assessed to determine the formation of the active site disulfide bond. [20]

2.7.2 Metal ion oligomerisation

The protein sample (40 μM) was transferred into metal association buffer (Table 2.5) using a desalting column. Ni^{2+} was spiked in at a ratio of 1:6, 1:3, 3:1, 6:1 Ni^{2+}: Protein. This was left O/N at 4 ^{o}C. The sample concentration was checked prior to any further testing to ensure no protein had precipitated. The buffer exchange (removal of imidazole) was achieved via stepwise dialfiltration – 1:10 x dilution into coordination buffer, concentrate to a 1/10 of the volume (Centrifugal Concentrators Vivaspin 20 10K Cut Off**)** and repeat five times. The concentration was checked again and samples that showed significant precipitation where re-suspended in 10 mM EDTA but no further testing was done. Samples with a stable concentration were analysed on SEC and AUC (Chapter Two at 2.5.1 and 2.5.3). The structure of any large oligomers was identified using TEM.

The 1:1 metal to protein ratio showed the optimal oligomerisation conditions. Different divalent metals were also examined (Mg^{2+} and Co^{2+}) and the extent of assembly was analysed using SEC. The oligomers were disassembled by adding 50 mM EDTA. The dissociation was confirmed using SEC (Chapter Two at 2.5.1). Different buffers and pH were trialled in accordance to the protocol from Ardini *et al*.[4] The protein was transferred into pH 7.6 metal association buffer (Table 2.1.7.4). The protein was analysed on AUC prior to metal coordinatiton to take note of any differences that occurred because of the change in buffer conditions. The coordintation with metal proceeded as above and the oligomers was analysed on AUC. The structure was visualised using TEM

2.7.3 pH switch

The protein (1 mg/mL) was exchanged into the varying pH buffers (Table 2.6) by O/N dialysis using either 2 mL dialysis buttons (Tube-O-DIALYZER, medi, 8K MWCO, Bioscience) or 100 µL dialysis buttons (Slide-A-Lyzer, MINI Dialysis Device, 10K MWCO, Thermo Fisher Scientific) depending on the amount needed for analysis. The dialysis was carried out at 4 ^{o}C excluding the DLS control assays in which samples were dialysed at room temperature. The pH of the sample was checked after dialysis to confirm successful buffer exchange. The samples were filtered through a 0.5 mL, 0.2 µm Eppendorf spin filter and the concentration was checked using either Bradford assay or Nanodrop to ensure no protein had precipitated. The protein oligomeric state was primarily analysed using SEC (Chapter Two at 2.5.1). Further solution analysis was conducted using AUC and DLS (Chapter Two at 2.5.3 and2.5.5). The structure of the oligomers was determined by TEM (Chapter Two at 2.5.4).

2.7.4 Protein concentration assay

1 mL of protein sample was exchanged into the different pH buffers as described above (pH 7.2, 7.6 and 7.8) at varying concentrations (0.5, 1, 3 and 5 mg/mL). The sample was filtered and both the concentration and pH were checked prior to further analysis. The molecular weight was analysed using HPLC with a static light scattering detector.

2.7.5 High salt

Non-reduced *Hs*Prx3at pH 8.0 (Table 2.5) was subjected to a range of different salt concentrations to disrupt the electrostatic interaction that were aiding the assembly of dodecamers into stacks and tubes.

2.7.5.1 NaCl or $(NH_4)_2SO_4$

Final salt conc (mM)	Volume of protein sample (µL)	Volume of concentrated salt buffer (µL)	Volume of no salt buffer (µL)
0	100	0	100
100	100	25	75
200	100	50	50
400	100	100	0

Table 2.18

The sample (2 mg/mL) was transferred into a no salt buffer (20 mM HEPES, pH 7.4) by O/N dialysis (Tube-O-DIALYZER, medi, 8K MWCO, Bioscience) at 4 ^{o}C. A (Table 2.6). The salt concentration was increased by diluting the sample at different ratios of high salt A or B buffers (Table 2.6, NaCl and $(NH_4)_2SO_4$ respectively) so the final protein concentration was 1 mg/mL and salt concentration varied (Table 2.18). The protein sample was incubated at 4 ^{o}C and concentration was analysed using Bradford assay at time points 30 mins, 60 mins, 90 mins, 180 mins and 21 hours. Before concentration analysis the sample was centrifuged to remove any precipitated protein.

2.7.5.2 Disassembly of tubes at pH 7.4

NaCl was use during this assay due to increased protein precipitation when using $(NH_4)_2SO_4$. The protein was dialysed O/N (Slide-A-Lyzer, MINI Dialysis Device, 10K MWCO, Thermo Fisher Scientific, 4 ^{o}C) into pH 7.4 buffer (Table 2.6) and the sample was analysed using SEC to ensure full oligomerisation. The protein solution was split into 3 x 0.5 mL and dialysed O/N into the different salt concentration buffers (High salt C, D and E, table 2.6). At this point the concentration was checked by Bradford assay and the pH was measured to ensure no variation due to buffer exchange. The sample was analysed using SEC and the MW calculated from the calibration curve (Chapter Two at 2.5.1).

2.7.6 Peroxiredoxin activity assay in competition with horse radish peroxidase

Buffers and solutions	
Activity assay buffer	100 mM KH_2PO_4, O.1 DPTA, pH 7.5
1M β-mercaptoethanol	15 µL 14.35 M stock, 85 µL uH_2O, pH 7 - 8
H_2O_2 6 x 10 $^{-4}$ % w/v	serial dilutions from stock (6 % w/v) - 1:10, 1:100, 1:1000, 1:10000 (1 mL)
HRP solution (~ 80 µM)	5 mg HRP-VI, 1 mL assay buffer

Table 19*All solutions were made in uH_2O and pH was adjusted using 1 M KOH in uH_2O.*

2.7.6.1 Sample preparation

*Hs*Prx3 was concentrated to 60 µL (between 15 – 20 mg/mL) using a Centrifugal Vivaspin 500 (µL) 10K cut off concentrator. 2 µL of 1 M β-mercaptoethanol was added to the sample and the mix was incubated on ice for 30 mins – 1 hour to ensure complete reduction of the peroxidatic cysteine. [21] The protein was loaded onto a Bio-Rad Micro-Spin Chromatography column that was primed with the assay buffer. To exchange the assay buffer and removed β-mercaptoethanol, the column was spun for 10 mins, 10, 000 x *g* at 4 ^{o}C and the flow through collected (repeat 4 x for > 99.5% buffer exchange). The protein concentration was measured prior to any analysis (Chapter Two at 2.4.1.2).

H_2O_2 substrate was prepared to a final concentration of 6 x 10 $^{-4}$ % w/v in ultrapure water (uH_2O) and the exact concentration was calculated from the absorption at 240 nm of a 6 x 10 $^{-3}$ % w/v dilution.

$$[H_2O_2\ 6 \times 10^{-4}\ \%\ w/v] = \frac{Abs\ 240}{\varepsilon 240 \times 10}$$

$$\varepsilon_{240} = 43.6\ M^{-1}\ cm^{-1}$$

Abs_{240} = absorbance of 6 x 10 $^{-3}$ % w/v dilution

Equation 2.5

To prepare the horse radish peroxidase solution, 5 mg was added to 1 mL of assay buffer. The concentration was calculated from the absorption at 403 nm of a 1:10 dilution

$$[\text{HRP}] = \frac{Abs\ 403\ x\ 10}{\varepsilon 403}$$

$\varepsilon_{240} = 1.02 \times 10^5\ M^{-1}\ cm^{-1}$

Abs_{240} = absorbance of 1:10 dilution

Equation 2.6

2.7.6.2 Assay

For the activity assay, all protein concentrations were assessed in triplicate. The spectrophotometer (Spectra max M5) was allowed to equilibrate to 25 ^{o}C before starting the assay. Each reaction well contained, *Hs*Prx3 (0 μM – 10 μM), H_2O_2 (8 μM) and HRP (8 μM), with the reaction volume being adjusted to 250 μL per well with the addition of assay buffer. *Hs*Prx3 (varying concentration) and HRP (8 μM) were pre-incubated at room temperature for at least 10 minutes prior to starting the assay, and then pipetted, in triplicate, into each well of a microtitre plate (Greinier 96-well clear bottom plate, Sigma-Aldrige). H_2O_2 solution was added to the wells directly before commencing the assay. The absorbance at 403 nm was analysed over 2 minutes with one scan every 8 seconds.

2.7.6.3 Data analysis

The peroxidatic activity of *Hs*Prx3 was measured by determining the degree to which *Hs*Prx3 out competed HRP in catalysing the substrate. HRP changes in colour from orange to clear and therefore the catalytic inhibition was measure by monitoring the change in solution absorbance at 403 nm. k_{Prx} was calculated using equation 2.6.

$$\left(\frac{F}{1-F}\right) k_{HRP}[HRP] = k_{Prx}[Prx]$$

$k_{HRP} = 1.78 \times 10^{7} M^{-1}\ s^{-1}$ [22]

(F/(F-1))= ratio of inhibition by HRP

Equation 2.6

The ratio of inhibition was derived from the change in absorbance at 403 nm (ΔAbs_{403}) when compared to the HRP reaction with no *Hs*Prx3. The baseline for the reaction was recorded from the assay blank (no peroxidase) (fig 2.3).

A: ΔAbs_{403} **B: $\left(\frac{F}{1-F}\right) k_{HRP}[HRP]$ against [Prx]**

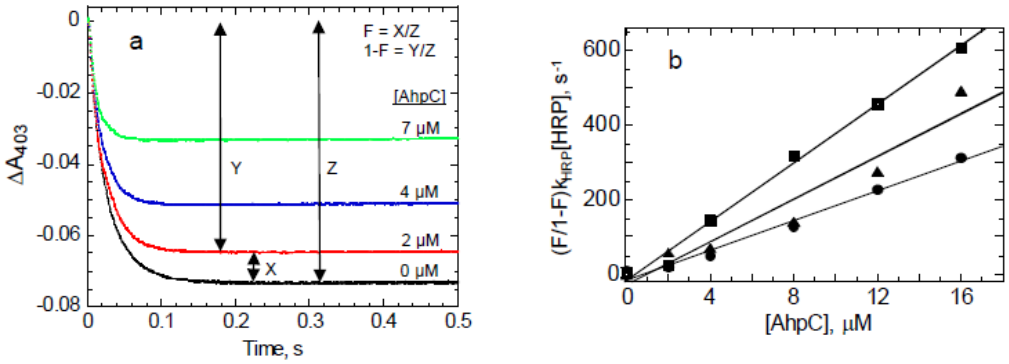


Figure 2.3: (A) Change in absorbance at 403 nm, black; just HRP, red, blue and green; HRP + increasing amounts of peroxiredoxin. (B) $\left(\frac{F}{1-F}\right) k_{HRP}[HRP]$ plotted against [Prx] to give k_{Prx}. [22]

$$F = x/z \qquad \textbf{Equation 2.7}$$

$$F\text{-}1 = y/z \qquad \textbf{Equation 2.8}$$

$\left(\frac{F}{1-F}\right) k_{HRP}[HRP]$ was plotted against [Prx] and the slope gave the second order reaction rate k_{Prx} with the units $M^{-1}s^{-1}$.

2.8 References

1. Bykowski T, Stevenson B: Aseptic technique. *Current protocols in microbiology* 2008, Appendix 4:4D-Appendix 4D.
2. Studier FW: Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* 2005, 41(1):207-234.
3. Good NE, Winget GD, Winter W, Connolly TN, Izawa S, Singh RMM: Hydrogen ion buffers for biological research. *Biochemistry* 1966, 5(2):467-&.
4. Ardini M, Giansanti F, Di Leandro L, Pitari G, Cimini A, Ottaviano L, Donarelli M, Santucci S, Angelucci F, Ippoliti R: Metal-induced self-assembly of Peroxiredoxin as a tool for sorting ultrasmall gold nanoparticles into one-dimensional clusters. *Nanoscale* 2014, 6(14):8052-8061.
5. Baneyx F: Recombinant protein expression in Escherichia coli. *Current Opinion in Biotechnology* 1999, 10(5):411-421.
6. Birnboim HC, Doly J: Rapid alkaline extraction proceedure for screening recombinant plasmid DNA. *Nucleic Acids Research* 1979, 7(6):1513-1523.
7. Summers EL, Meindl K, Uson I, Mitra AK, Radjainia M, Colangeli R, Alland D, Arcus VL: The structure of the oligomerization domain of Lsr2 from Mycobacterium tuberculosis reveals a mechanism for chromosome organization and protection. *PloS ONE* 2012, 7(6).
8. Duong-Ly KC, Gabelli SB: Gel filtration chromatography (size exclusion chromatography) of proteins. In: *Laboratory Methods in Enzymology: Protein, Pt C*. Edited by Lorsch J, vol. 541. San Diego: Elsevier Academic Press Inc; 2014: 105-114.
9. Kagedal L, Engstrom B, Ellegren H, Lieber AK, Lundstrom H, Skold A, Schenning M: Chemical, Physical and chromatographic properties of SUperddex-75 prep grade and Superdex-200 prep grade gel-filtration media. *Journal of Chromatography* 1991, 537(1-2):17-32.
10. Kunji ERS, Harding M, Butler PJG, Akamine P: Determination of the molecular mass and dimensions of membrane proteins by size exclusion chromatography. *Methods* 2008, 46(2):62-72.
11. Blommel PG, Fox BG: A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expression and Purification* 2007, 55(1):53-68.
12. Bradford MM: Rapid and sensitive method for quantification of microgram quantities of protein utilizing princle of protein-dya binding *Analytical Biochemistry* 1976, 72(1-2):248-254.
13. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A: ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* 2003, 31(13):3784-3788.
14. Shapiro AL, Vinuela E, Maizel JV: Molecular weigh estimation of polypeptide chaine by electrophoresis in SDS-polyacrylamide gels. *Biochemical and Biophysical Research Communications* 1967, 28(5):815-&.
15. Ghaemmaghami S, Huh W, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS: Global analysis of protein expression in yeast. *Nature* 2003, 425(6959):737-741.
16. Greenfield NJ: Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* 2006, 1(6):2876-2890.
17. Cole JL, Lary JW, Moody T, Laue TM: Analytical Ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods in Cell Biology* 2008, 84:143-179.

18. Schuck P: Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophysical Journal* 2000, 78(3):1606-1619.
19. Carvalho JWP, Santiago PS, Batista T, Garrido Salmon CE, Barbosa LRS, Itri R, Tabak M: On the temperature stability of extracellular hemoglobin of Glossoscolex paulistus, at different oxidation states: SAXS and DLS studies. *Biophysical Chemistry* 2012, 163:44-55.
20. Gourlay LJ, Bhella D, Kelly SM, Price NC, Lindsay JG: Structure-function analysis of recombinant substrate protein 22 kDa (SP-22) - A mitochondrial 2-Cys peroxiredoxin organized as a decameric toroid. *Journal of Biological Chemistry* 2003, 278(35):32631-32637.
21. Hall A, Nelson K, Poole LB, Karplus PA: Structure-based insights into the catalytic power and conformational dexterity of Peroxiredoxins. *Antioxidants & Redox Signaling* 2011, 15(3):795-815.
22. Dolman D, Newell GA, Thurlow MD, Dunford HB: Kinetic study of reaction of Horseradish-Peroxidase with hydrogen-peroxide. *Canadian Journal of Biochemistry* 1975, 53(5):495-501.

Chapter Three: Controlled oligomerisation of Lsr2 and its derivatives

3.1 Introduction

Proteins are attractive building blocks (tectons) for the formation of biological nanomaterials due to their controllable assembly and the wealth of potential surface functionalisation sites that they provide. [1, 2] The self-association of protein subunits is intrinsically linked to the amino acid sequences at the interfaces, which have been arranged, by nature, to drive the formation of materials of a specific size, shape and function (Chapter One: at 1.3.3). [3, 4] This is inherently difficult to achieve using purely synthetic molecules due to the different conformations the self-assembling molecules can take, and the propensity of small (nanoscale) units, with a large surface area, to aggregate. [5] When identifying a protein as a tecton it is important to consider a number of characteristics:

(1) Is the protein readily available?
(2) Can the self-assembly be controlled?
(3) Is there in-built function or the opportunity to incorporate function into the tectons?

Lsr2 has been selected in this study as it has the potential to encompass all of the criteria stated above. The N-terminal domain of the protein (Nterm-Lsr2) will also be analysed as it has been shown to undergo self-assembly in response to a specific enzymatic cleavage event. [6, 7]

Lsr2 is a small, basic DNA binding protein present in *Mycobacteria* that regulates gene expression and influences the organisation of the bacterial nucleoid (Chapter One at 1.5.1). [6, 8-10] It belongs to a group of proteins known as histone-like proteins that have similar functions to their eukaryotic counterparts (histone proteins). [11, 12] Although these proteins appear to function in a similar manner, non-specific binding to adenine and thymine (AT) rich regions of DNA, they have little homology in terms of sequence and structure. [11, 13, 14] Lsr2 has two domains that are both essential for function; the N-terminal oligomerisation domain and the C-terminal DNA binding domain. [6, 13] The DNA binding domain is comprised of two α-helices that are joined together by a disordered loop (Chapter One at

1.5.2). [14] It is unique to other histone-like proteins in that the α-helices are organised almost perpendicular to each other, stabilised in this conformation via a series of hydrophobic interactions from complementary residues on each α-helix. [14] The residues Arg^{97}-Glu^{98}-Arg^{99}, within the loop, bind to the minor groove of the DNA helix in a hook motif with the binding being stabilised by a hydrogen bonding network. [13, 14] It has been shown that the DNA binding efficiency is significantly diminished when Arg^{97} and Arg^{99} are substituted for an alanine, indicating that these residues are essential for function. [14] Lsr2 has the ability to form highly ordered, stable protein complexes. [10] Furthermore, the protein has been shown to organise DNA by bridging two DNA strands through the N-terminal dimerisation domain, [15] and even driving the formation of rigid nucleoprotein fibres. [12]

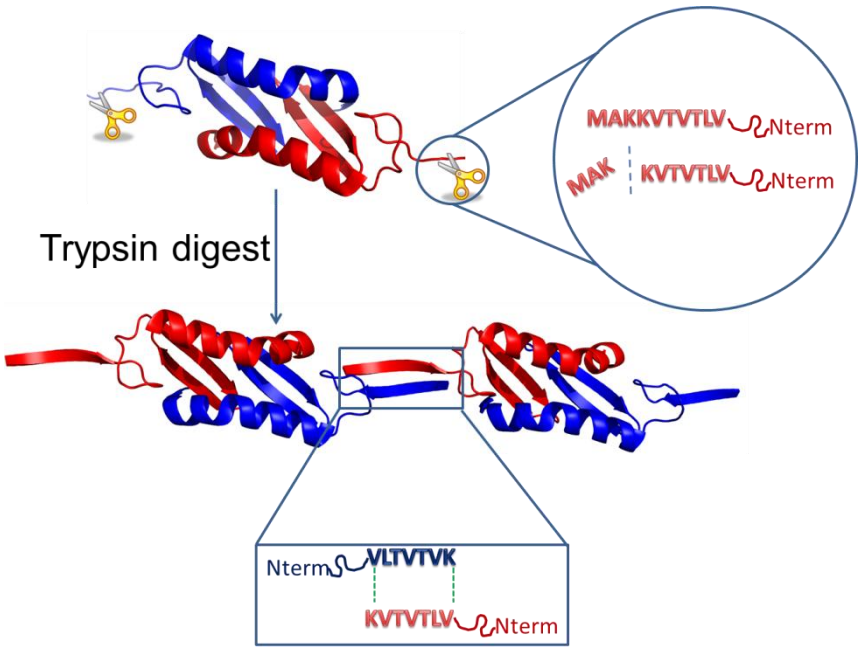


Figure 3.1: Pymol image of the oligomerisation via proteolytic cleavage of Nterm-Lsr2 [6] Removal of the N-terminal lysine removes electrostatic repulsion between the N-terminal extensions. This allows the formation of an inter-dimer β-strand between the hydrophobic residues in the random coil. *Images generated using Pymol with coordinates from PDB:4E1R.*

In a study by Summers and co-workers, the truncated form of Lsr2 (Nterm-Lsr2) was shown to oligomerise into linear chains via anti-parallel β-sheet formation between the extended N-

termini on neighbouring dimers (fig. 3.1) (Chapter One at 1.5.2). [6] This process has been achieved *in vitro* using trace amounts of trypsin enabling oligomerisation on-demand. [7] The protocol provides a useful tool for studying the onset of the oligomerisation of Nterm-Lsr2 and its uses as a protein tecton.

It is well established that changes in the physical environment of a protein can alter the structure, solubility, activity, and stability of a protein. [16-18] For example, the silk protein from *Bombyx mori* can form different quaternary structures in response to changes in pH. [19] To explore the full potential of the Nterm-Lsr2 tecton, impact of factors such as pH and protein concentration were also explored as alternate triggers to enzymatic cleavage.

By studying the full length protein and therefore combing the oligomerisation switch with the DNA binding efficiency, [9, 12, 14] it may be possible to design a protein tecton that can organise DNA in a controlled manner. Lsr2 has the potential to be utilised to form hybrid nucleoprotein materials which makes it an ideal protein for the discovery and development of a novel biological tecton.

3.2 Lsr2 as a protein tecton in bionanotechnology

Lsr2 has a number of potential functions that make it an attractive target for development into a protein tecton. In addition to the DNA binding domain there is also a poly-6-histidine tag, located at the C-terminus of the protein. Metal coordination sites can be used to direct assembly though the highly specific ligand-metal association geometry which, in turn, aids control over oligomerisation. [20-22] Metal-binding tags have also been utilised to anchor additives to protein nanostructures, for example, Ni^{2+} treated gold nanoparticles (AuNPs), bound through the strong Ni^{2+} to imidazole interaction. [23] This particular functionality is not available in the truncated protein (Nterm-Lsr2) due to the positioning of the tag at the N-terminus of the protein which is removed during proteolysis. [6, 7]

It is clear that Lsr2 satisfies at least one of the characteristics of a good protein tecton in that it has a wealth of in-built functions. The purpose of this research is to explore the other two characteristics; to express and purify Lsr2 in a recombinant protein expression system, and to instigate assembly in a controllable manner.

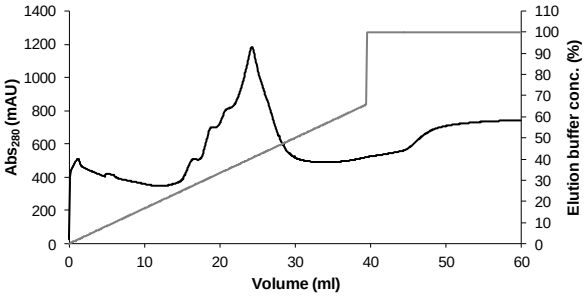
3.2.1 Protein expression and purification

The Lsr2 plasmid was kindly provided by Dr Vic Arcus at Waikato University. The gene had been cloned into the expression vector pET30b with a C-terminal histidine tag. Electrocompetent *E. coli* cells (BL21(DE3)*) were transformed with the plasmid using electroporation (Chapter Two at 2.2.2.1). The protein was successfully expressed in *ZYM-5052* auto-induction media [24] (Chapter Two at 2.1.6.2) in the presence of kanamycin after shaking at 180 rpm, at 37^{o}C for 4 hours, and then at 30^{o}C for a further 16 hours.

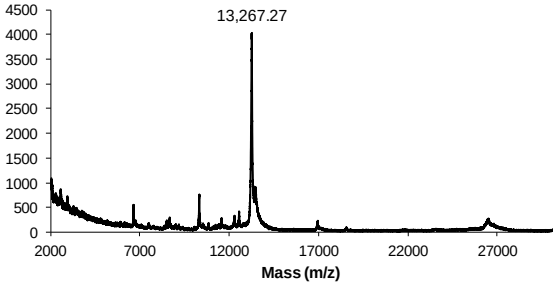
Cells were lysed using sonication (Chapter One 2.3.2.2) and the soluble and insoluble fractions were separated via centrifugation (20 mins, 18,000 x *g*). Lsr2 was isolated from the soluble crude fraction using immobilised metal-ion affinity chromatography (IMAC) (fig. 3.2 A) (Chapter Two 2.3.2.5) and the excess imidazole was removed by buffer exchange using a desalting column. Matrix-assisted laser desorption/ionisation time of flight mass spectroscopy (MALDI-TOF MS) showed the Lsr2 monomer molecular weight to be 13.26 kDa, slightly larger than the calculated molecular weight of 13.16 kDa (fig.3.2 B). The presence of Lsr2 in the eluted IMAC peak was confirmed by SDS-PAGE (fig. 3.2 C) (Chapter Two at 2.4.3). An additional band was also seen on SDS-PAGE. Initially it was believed that this could be a SDS stable oligomer, as had been characterised in previous studies. However, Western-blot

using the anti-his antibody, confirmed that this band did not relate to Lsr2 but was an impurity that did not contain a histidine tag (Appendix 3.1).

A: IMAC



B: MALDI-TOF MS



C: SDS-PAGE

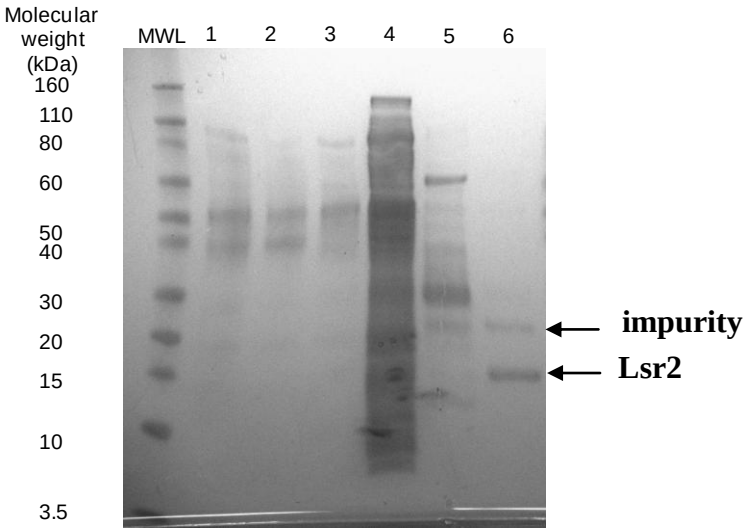


Figure 3.2: Purification of Lsr2. (A) Elution of Lsr2 from a Ni^{2+} IMAC column upon increasing imidazole concentration. Black: absorbance at 280 nm (mAU); grey: percentage concentration of 1 M imidazole buffer (%). (B) MALDI-TOF MS of purified Lsr2. The molecular weight (MW) is slightly larger than the calculated value of 13.16. (C) SDS-PAGE gel of crude fractions and purified Lsr2 from IMAC. 1: crude lysate; 2: insoluble crude; 3: soluble crude; 4: IMAC flow through; 5: impurity from IMAC; 6: isolated Lsr2 monomer + impurity.

Throughout the purification process there were significant amounts of impurities eluting with the target protein even in the presence of high salt (0.5 M). It is believed that the impurity

was either protein weakly bound to the IMAC column (high native histidine content) or trapped within a genomic DNA-Lsr2 complex. If the protein was weakly bound to the IMAC column then using a stepwise gradient to increase the imidazole elution buffer concentration in increments (0.02 – 1 M in a stepwise fashion) would aid the separation of eluted proteins into individual fractions (fig. 3.3 A). Several peaks eluted and the proteins within the peaks were identified on SDS-PAGE (fig. 3.3 B). The impurity was observed in the initial eluted fractions and only fractions containing isolated Lsr2 were collected for further analysis.

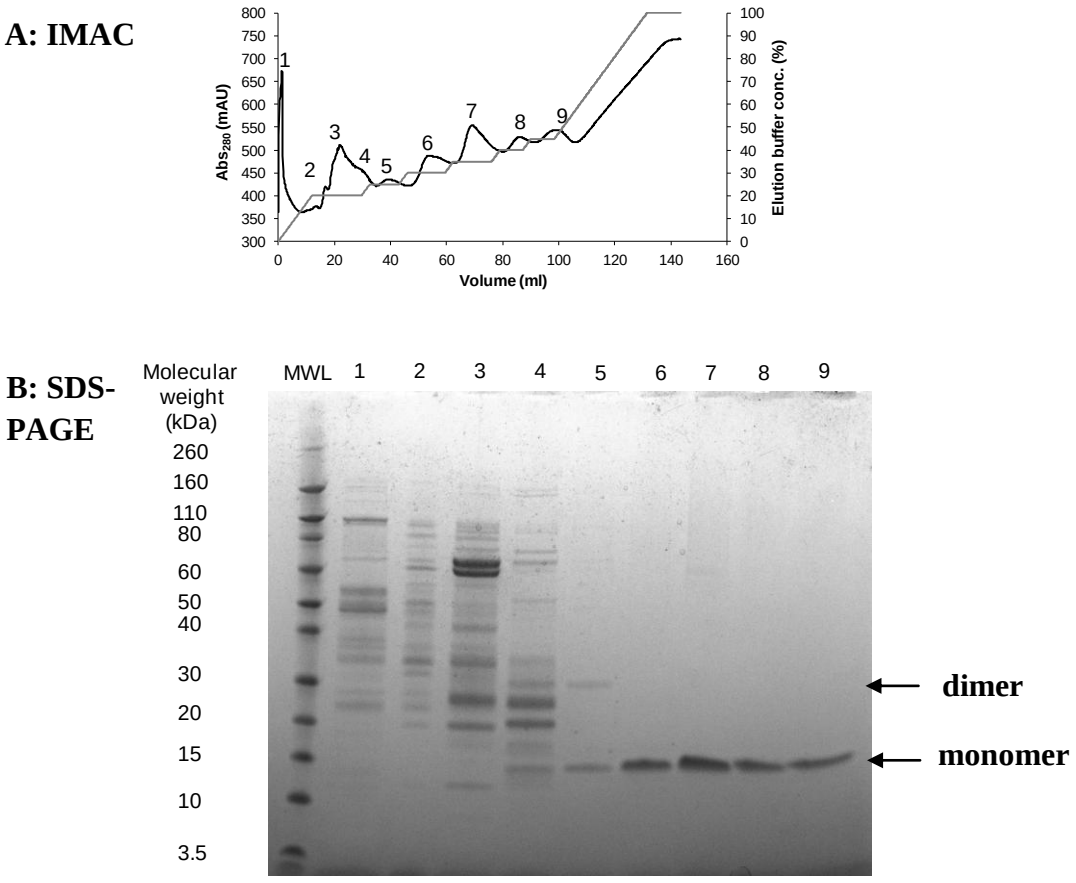


Figure 3.3: Stepwise purification of Lsr2 using a slow imidazole gradient. (A) Elution of Lsr2 and impurities from a Ni^{2+} IMAC column upon increasing imidazole concentration in a stepwise gradient. Black: Absorbance at 280 nm (mAU); grey: percentage concentration of 1 M imidazole buffer (%). Peaks 1 – 9 were analysed on SDS-PAGE. (B) SDS-PAGE gel of the fractions eluted on IMAC. Fractions 1-4: impurities; 5: Lsr2 monomer and dimer; 6 – 9: isolated Lsr2 monomer. Using a stepwise gradient allows for the elution of pure Lsr2 directly from the IMAC.

3.2.2 Trypsin digest of full length Lsr2

During a study by Summer *et al.* it was noted that the eluted protein from the IMAC was bound to a significant amount of genomic DNA. [6] Instead of removing the DNA completely, they utilised the bound DNA to study the effects of DNA compaction and protection by Lsr2 in response to trace amounts of trypsin and DNAse. The protein that had not been treated with trypsin was unable to protect the genomic DNA from digest, as visualised by DNA agarose gel. [6] However, treatment of the protein-DNA sample with trypsin for increasing amounts of time not only showed that the DNA was compacted with increasing amounts of proteolysis time but, more intriguingly, that it was protected from digestion by DNAse. [6] Furthermore, the band corresponding to the monomer completely disappeared on SDS-PAGE with increasing amounts of tryspin digest time but the bands relating to the dimer and HMW oligomers did not. This further strengthened the argument that Lsr2 is used in nature to protect DNA and it does this by oligomerisation at the N-terminal domain to form HMW protein oligomers that are resistant to denaturation by heat and SDS.

The experiments were repeated in this study to ensure the protein was still able to undergo proteolysis with trypsin. Lsr2 (1 mg/mL) was subjected to a trypsin digest (1:500, trypsin:Lsr2 w/v) for increasing amounts of time (5, 10, 15, and 30 mins) at room temperature. The digest was stopped via heat denaturation of the protein and enzyme in a water bath set at 95^{o}C (5 mins). The denatured sample was then split for both SDS-PAGE and DNA agarose gel analysis (fig. 3.4 A and B respectively) with trypsin run as a control. Unlike the previous study, the monomer band was still prevalent on SDS-PAGE, even after 30 minutes although it had faded slightly. Interestingly, the protein, which had previously appeared as a single monomer band by SDS-PAGE, had assembled into SDS stable oligomers while being stored at 4^{o}C. It is possible that the degree of DNA binding had increased over time, stabilising the complex even under protein denaturing conditions. The HMW and dimer bands also showed little variability on SDS-PAGE. The DNA agarose gel displayed that the Lsr2 had indeed eluted with significant amounts of DNA with all of the lanes containing Lsr2 showing fluorescence with DNA loading dye. The only lane that had no fluorescence contained trypsin only, showing that protein alone was not enough to appear under UV light. While there was a small amount of compaction of the DNA band it was not to the same extent as noted in the literature. [6]

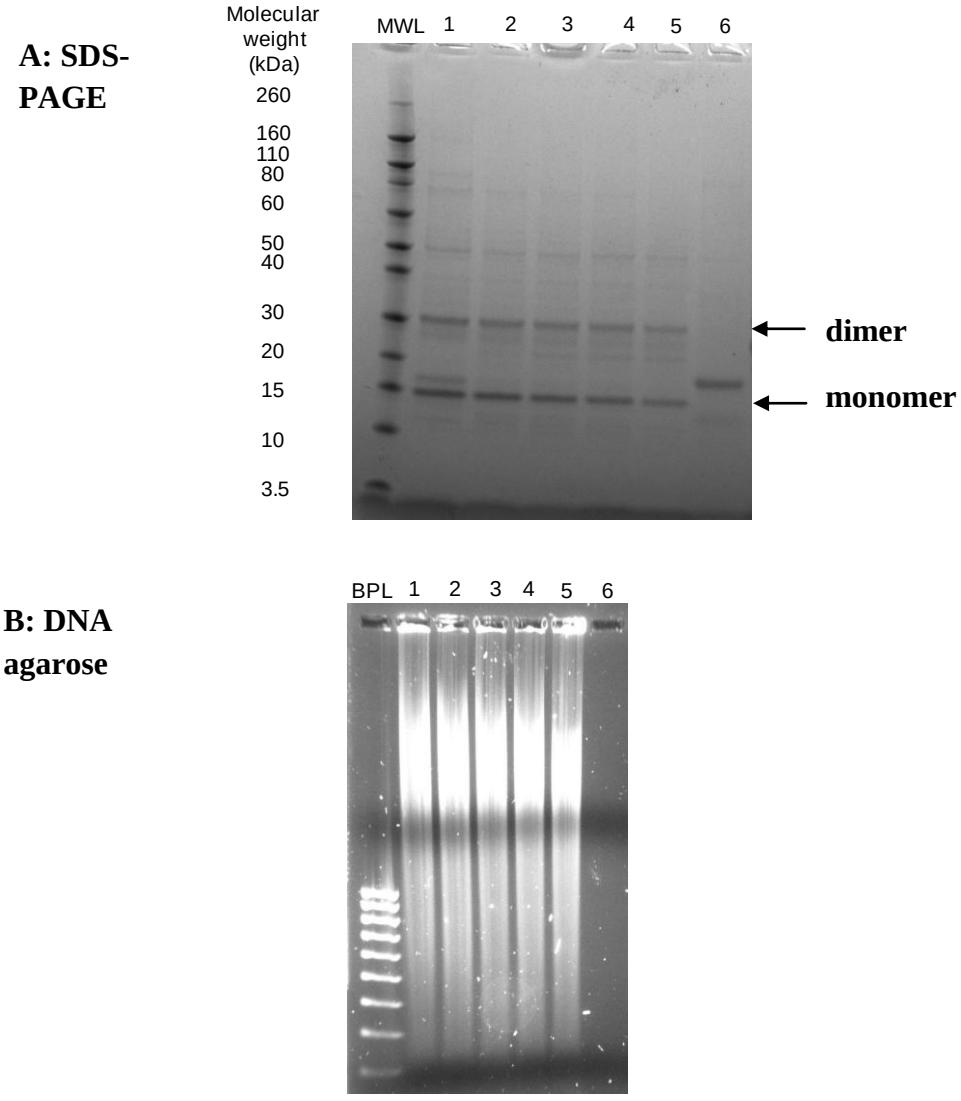


Figure 3.4: Electrophoresis analysis of trypsin proteolysis. (A) SDS-PAGE gel of trypsin proteolysis and trypsin. 1: Lsr2 no trypsin; 2, 3, 4 and 5: Trypsin + Lsr2 (1:500) proteolysis for 5, 10, 15, and 30 minutes respectively; 6: trypsin. (B) DNA agarose gel of trypsin proteolysis and trypsin as a negative control. 1: Lsr2 no trypsin; 2, 3, 4 and 5: trypsin + Lsr2 (1:500) proteolysis for 5, 10, 15, and 30 minutes respectively; 6: trypsin. All Lsr2 samples contain bound DNA. This is possibly protecting the protein from trypsin digest as very little degradation is seen on SDS-PAGE.

While this experiment provided a simple means to study potential DNA compaction in response to Lsr2 oligomerisation there were too many variables and therefore it was not possible to repeat the protocol using exactly the same conditions. The main issue is that there

is no control over the concentration of bound DNA. It is possible that the amount of bound DNA in the current study exceeded the amounts characterised by Summers *et al.* This would explain the lack of trypsin digest as cleavage sites may be blocked by the high concentration of DNA. It would also explain the decrease in compaction, as the ratio of protein to DNA is less and therefore the DNA compaction is less. To accurately characterise the DNA compaction it is necessary to remove the genomic DNA and then replace it with DNA of a known concentration.

3.2.3 Strong DNA binding of Lsr2

The denatured form of Lsr2 was a monomer at around 13 kDa. This was seen both on SDS-PAGE and MALDI-TOF MS (fig. 3.2, B and C above). However, when the sample was run on size exclusion chromatography (SEC) the major peak eluted at the void which indicated an oligomer greater than 660 kDa (fig. 3.4 A). SDS-PAGE molecular weight analysis showed that while the other small peaks seen on the SEC trace contained mainly small amounts of impurities, Lsr2 eluted almost entirely at the column void (fig. 3.4 B).

As discussed above, Lsr2 was bound to large amounts of genomic DNA. This has been seen in previous studies using techniques such as AFM and TEM which showed that Lsr2 has the propensity to form large complexes with DNA. [6, 12, 15] The DNA binding is likely to have occurred during cell lysis where large amounts of genomic DNA are released from the cell. To accurately assess any oligomerisation that was a direct consequence of trypsin digest and therefore probe the capabilities of Lsr2 as a protein tecton, it was necessary to try and remove this DNA.

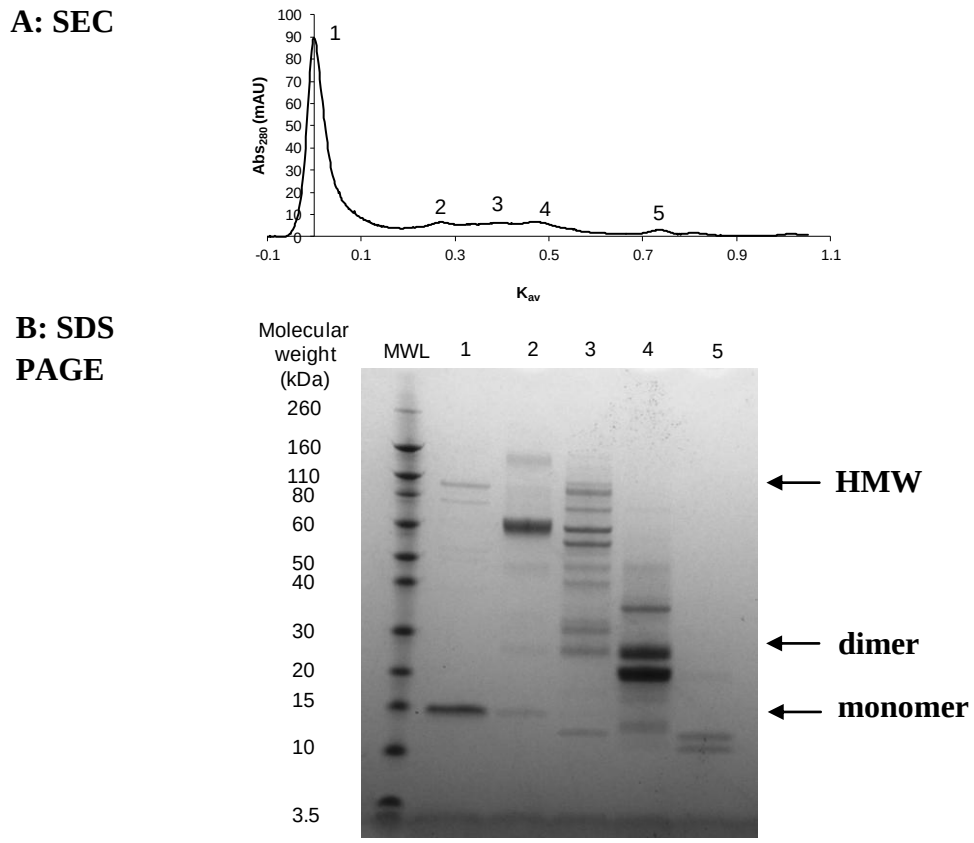


Figure 3.5: Size exclusion chromatography (SEC) analysis of Lsr2. (A) SEC of Lsr2 from IMAC (50 mM sodium phosphate, 200 mM NaCl, pH 7.4). (B) SDS-PAGE gel of the SEC peaks 1- 5. Almost all of the Lsr2 has eluted at the void volume (> 660 kDa), with the peaks at MW containing impurities. Lsr2 exists as a dimer at 26 kDa so the high MW elution volume indicates that the protein has aggregated.

DNA binding disruption by environmental conditions

Numerous methods were employed to try and separate Lsr2 from DNA. The primary trials involved simple environmental conditions such as adding a high concentration of salt and the destruction of the DNA via digest with DNAse. These simple variations were not sufficient to completely remove the DNA. Agarose gel electrophoresis showed that there was still a significant amount of intact DNA in the protein sample and SEC analysis of the high salt concentration still displayed a large peak around the column void (Appendix 3.2). A combination of long range, complex and non-specific interactions between the DNA and Lsr2

could be reducing the effect of high salt and helping Lsr2 stay bound to DNA. Furthermore, the DNA is likely to be protected from digest by DNAse while still bound to Lsr2 as this protein has been identified as protecting DNA from attack from outside sources. [6, 11]

DNA binding disruption by site specific substitution

The NMR structure of the C-terminal region of Lsr2 displays two residues that appear to be essential for DNA binding. [14] These are Arg^{97} and Arg^{99}, located in the connecting loop region and involved in the DNA binding hook motif (fig. 3.6). The proton signals for these residues completely disappear in the presence of DNA. [14] This is indicative of proton transfer between the DNA and protein on the NMR time scale, an indication of hydrogen bonding. [25] Furthermore, substitutions of these residues for alanine have been shown to produce constructs with significantly diminished DNA binding efficiency when compared to the wild type. [13] For this reason, the mutants were employed to try and remove the DNA from the full length Lsr2 construct.

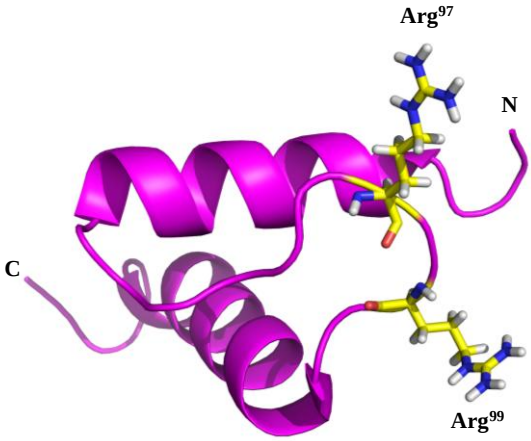


Figure 3.6: NMR structure of the C-terminal DNA binding domain. [14] The DNA-binding residues are highlighted in yellow with hydrogen – white, nitrogen – blue, and oxygen – red. These residues were sequentially mutated to alanine to reduce the DNA binding efficiency. *Images generated using Pymol with coordinates from PDB:2KNG.* [14]

The plasmids encoding the mutated proteins (Lsr2-R97A, Lsr2-R99A, and Lsr2-R97AR99A) were supplied by the Arcus lab at The University of Waikato.

Mutant	Site specific substitution
Single mutant 1	Lsr2-R97A
Single mutant 2	Lsr2-R99A
Double mutant	Lsr2-R97AR99A

Table 3.1: The site specific mutations of Lsr2. The proposed DNA binding residues are removed to try and disrupt the strong Lsr2-genomic DNA interaction.

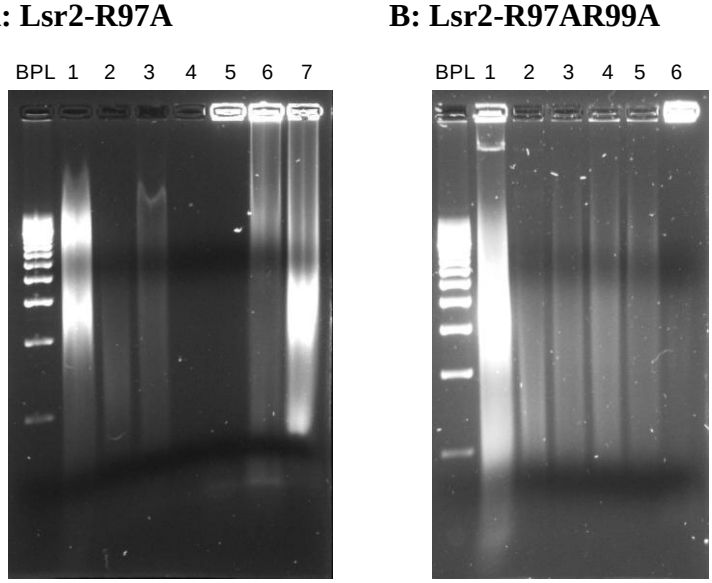


Figure 3.7: Agarose gel analysis of DNA binding arginine mutants. (A) Lsr2-R97A mutant agarose gel, 1 – crude lysate, 2 – insoluble crude, 3 – soluble crude, 4 – IMAC flow through, 5 – IMAC peak 1, 6 – IMAC peak 2, 7 – IMAC peak 3 (all IMAC peaks contain Lsr2-R97A). (B) Lsr2-R97AR99A mutant agarose gel, 1 – crude lysate, 2 – insoluble crude, 3 – soluble crude, 4 – IMAC flow through, 5 – IMAC peak 1, 6 – IMAC peak 2 (all IMAC peaks contain Lsr2-R97AR99A).

One single mutant, Lsr2-R97A and the double mutant were selected for analysis. Both were expressed and purified using the methods from section 3.2.1 (above). The eluted fractions from the IMAC were analysed on SDS-PAGE and agarose gel electrophoresis.

For both constructs the denatured protein showed as a single monomer band on SDS-PAGE (Appendix 3.3). However, DNA agarose gel showed that the fractions from the IMAC that

were related to the purified protein still contained a significant amount of genomic DNA (fig. 3.7 A and B).

Interestingly, the bound DNA is unable to travel through the gel with the fluorescence relating to the DNA-protein final eluted IMAC peak being entirely within the well. This indicates large, complex structures that are too big to be transported through the agarose gel.

3.2.4 Discussion

While the Lsr2 was successfully produced from a readily available source, all attempts to isolate the protein from contaminating genomic DNA were ineffective. In all conditions, including site specific substitutions, the DNA was still tightly bound to the protein. This is possibly due to the high concentration of DNA released during cell lysis, and a large amount of non-specific biding by the full length protein.

For a protein to be considered as a tecton for nanotechnology there needs to be a degree of control over the assembly. Analysis of the controlled oligomerisation via proteolytic cleavage would be unreliable while Lsr2 was still bound to DNA, as it would be difficult to separate out the different routes towards assembly. In addition, it would be challenging to employ specific DNA binding to direct assembly of nucleoprotein complexes while the protein was still bound in a non-specific manner to genomic DNA. Overall, it was deemed that the use of Lsr2 as a protein tecton was not possible with the current methodology. For this reason the focus was directed towards controlling and improving upon the protocol for assembling the truncated form, Nterm-Lsr2, which removed the DNA binding site, thereby eliminating the complications associated with this particular function.

3.3 Nterm-Lsr2 oligomerisation

The N-terminal dimerisation domain of Lsr2 has been identified as the oligomerisation region of the protein and is essential for the continued function of the protein (Chapter 1 at 1.5.2). Lsr2 has the ability to bind to DNA, bridging non-specific sequences and even forming elongated nucleoprotein fibres with both single-stranded and double helix DNA. [12, 15] These attributes, as well as the existence of SDS resistant HMW oligomers seen on SDS-PAGE, suggest that the protein is able to oligomerise into higher ordered structures, with the N-terminal dimerisation domain facilitating this interaction.

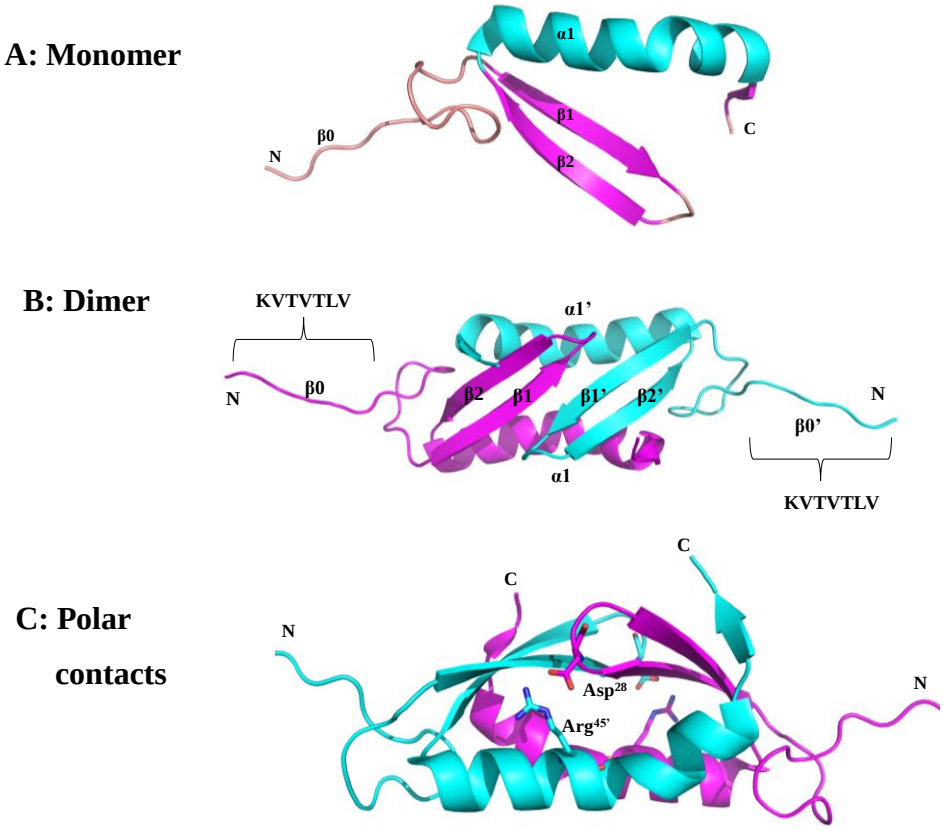


Figure 3.8: Crystal structure of the N-terminal oligomerisation domain. (A) Nterm-Lsr2 monomer, one anti-parallel strand, one random coil and one α-helix. (B) Nterm-Lsr2 monomer which forms via an anti-parallel β-strand interaction between β1 and β1' to form a four stranded β-sheet which buries ~ 31 % surface area per monomer. (C) The polar contacts that further stabilise the dimer interface. *Images generated using Pymol with coordinates from PDB: 4E1P*[6]

The N-terminal domain monomer (residues 1 – 61) is a random coil at the N-terminus followed by a β-strand and a single α-helix with a significant kink close to the C-terminus (fig. 3.8 A). The β-strand and α-helix axis are arranged at approximately a 45^{o} angle with respect to each other. Dimerisation occurs between β1 and β1', forming a four stranded β-sheet that buries around 31 % of the monomer surface (~1548 $Å^2$) (fig. 3.8 B). The kink in the α-helix allows for compact packing against the adjacent β-strand and the interface is further stabilised by electrostatic interactions between Asp^{28} and $Arg^{45'}$ from α1' (fig. 3.8 C). The N-terminal residues are remarkably well suited for β-strand formation, comprised almost entirely of hydrophobic residues (fig. 3.8 B).

When the crystal structure was solved it was noted that the dimers did indeed oligomerise via a previously uncharacterised inter-dimer β-strand interaction between the N-terminal random coils. [6] There was no discernible electron density for either the histidine tag or residues Met^1, Ala^2 and Lys^3. The mass of the crystallised protein, determined by electrospray mass spectroscopy, was 9198.0 m/z, close to the calculated value of an Nterm-Lsr2 monomer that had undergone cleavage between Lys^3 and Lys^4 (6199.8 Da). It was noted that crystals only grew after the protein had been left for some time to undergo an unknown maturation process. This process was mimicked by the addition of trace amounts of trypsin (1:500, trypsin:Nterm-Lsr2) which cleaved between the two N-terminal lysine residues. The structure determined from this process also showed linear chains of Nterm-Lsr2, joined together by β-strands at the inter-dimer interface. [6]

Trypsin cleavage at Lys^4 allows for the hydrophobic residues on the N-terminal random coils on neighbouring dimers to interact and form a β-strand. Without the removal of the N-terminal residues, electrostatic repulsion and steric clashes between these residues would hinder oligomerisation. In addition to the hydrophobic interactions, there is a complex network of polar interactions between the positively charged primary amine of Lys^4 and the carbonyl and acidic groups of the neighbouring dimer which further stabilises the inter-dimer interface. This would not be possible without the removal of the histidine tag and three N-terminal residues.

Results from the current study concur with Summers' findings and reinforce the idea that controlled proteolysis of the N-terminal residues facilitates the formation of stable HMW oligomers from the truncated construct of Lsr2 (Nterm-Lsr2).

3.3.1 Expression, purification and characterisation of Nterm-Lsr2

Nterm-Lsr2 was expressed in *E .coli* BL21(DE3) cells using standard procedures as described in section 3.2.1. It was purified from other soluble proteins in the crude lysate using IMAC (fig. 3.9 A) and SEC (fig. 3.9 B). Expression and purification of Nterm-Lsr2 was verified by SDS-PAGE analysis (fig. 3.9 C).

Circular dichroism (CD) spectroscopy showed that the protein was correctly folded with troughs at 209 and 219 nm representing the α-helix and β-sheet respectively (fig. 3.10 A). MALDI-TOF MS confirmed the molecular weight of the monomer to be 9470 Da (fig. 3.10 B). Although MALDI-TOF MS is considered to be a soft ionisation method, ideal for the analysis of biological molecules, [26] there have been cases where the ionisation is enough to fragment the C-terminal residue. [26] The MW of Nterm-Lsr2 is slightly lower that the calculated molecular weight of 9626 Da but this is consistent with the mass if the C-terminal arginine residue has been lost.

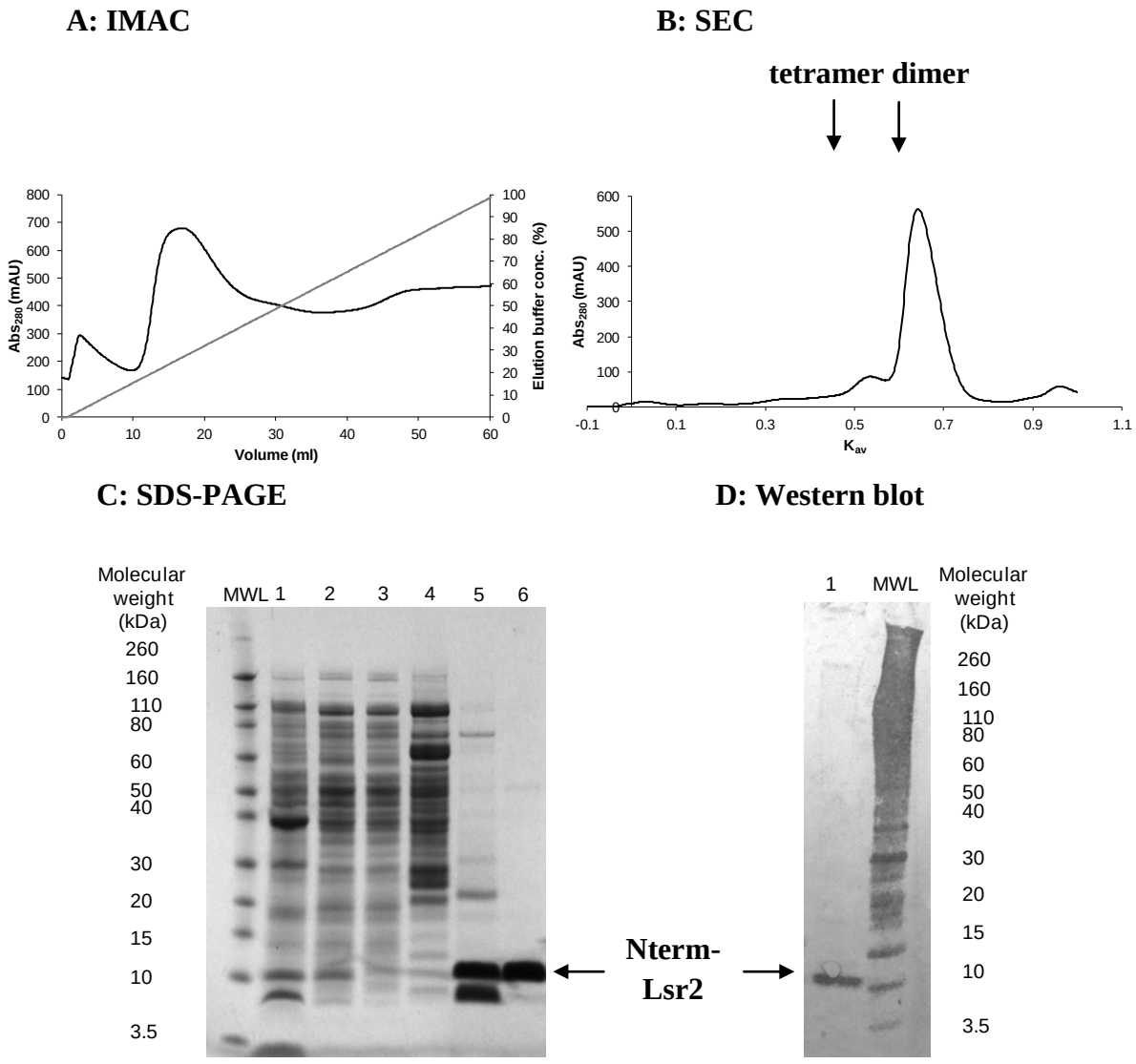


Figure 3.9: Purification and isolation of Nterm-Lsr2. (A) Elution of Nterm-Lsr2 from a Ni^{2+} IMAC column upon increasing imidazole concentration. Black: Absorbance at 280 nm (mAU); grey: percentage concentration of 1 M imidazole buffer (%). (B) SEC chromatography of eluted IMAC peak showing a prominent dimer peak (19 kDa) and a minor peak around the MW of a tetramer (40 kDa). This may be due to the high loading concentration (~10 mg/mL). (C) SDS-PAGE gel of protein extraction and purification. 1: crude lysate; 2: insoluble crude; 3: soluble crude; 4: IMAC flow through; 5: IMAC imidazole eluted fractions; 6: isolated Nterm-Lsr2 from SEC. (D) Western blot of Nterm-Lsr2. 1: Nterm-Lsr2 and 2: mouse anti-polyhistidine tagged ladder. Electrophoresis shows a histidine tagged protein at the MW of an Nterm-Lsr2 monomer.

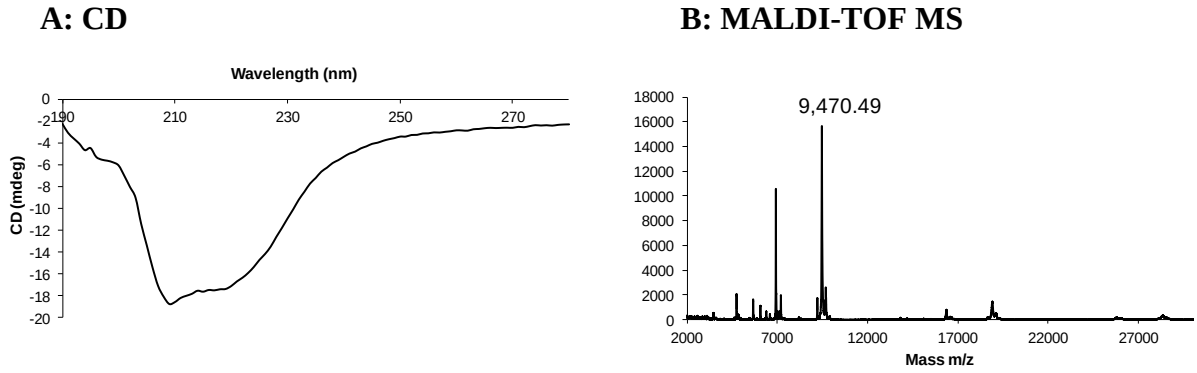


Figure 3.10: Characterisation of Nterm-lsr2 dimer. (A) CD of full length Nterm-Lsr2. The troughs at 209 and 219 nm show the secondary structural elements of α-helix and β-sheet respectively. (B) MALDI-TOF MS of full length Nterm-Lsr2, major peak at 9470 m/z which is consistent with the calculated monomer molecular weight if the C-terminal arginine has been lost.

3.3.2 Proteolytic oligomerisation of Nterm-Lsr2

SEC was used as a simple analytical tool to map the changes from non-oligomerised and oligomerised protein by looking at changes in MW. On SEC the non-oligomerised (neat sample) eluted at a volume consistent with the dimer MW (19 kDa) (fig. 3.11 B, peak 2), around double that of the monomer molecular weight from mass spectroscopy (9473 Da). Trypsin cleavage at Lys^4 not only removed Met^1, Ala^2, and Lys^3 but also the poly-6-his tag, TEV recognition site and linker which results in a 3444.7 Da (calculated) loss in the molecular weight of the monomer after proteolysis. The mass of the digested monomer, determined by MALDI-TOF MS, was 6188 Da (fig. 3.11 A). The dimer was also seen on MS with a signal at 12361 m/z.

By setting the dimer molecular weight at 12.4 kDa, the number of dimeric units in each oligomer could be approximated from the mass calculated from the SEC calibration curve. Upon partial digestion with trypsin (1:500 w/v) Nterm-Lsr2 oligomerised into three sets of HMW structures (fig. 3.11 B). The oligomers were calculated to be a dodecamer, a hexamer and a tetramer. The fourth peak on the spectrum was shown, through SDS-PAGE analysis, to be protein fragments (Appendix 3.4). Fragmentation of Nterm-Lsr2 is a result of trypsin proteolysis at the additional lysine and arginine residues along the polypeptide chain. This was seen in all partial trypsin digest experiments.

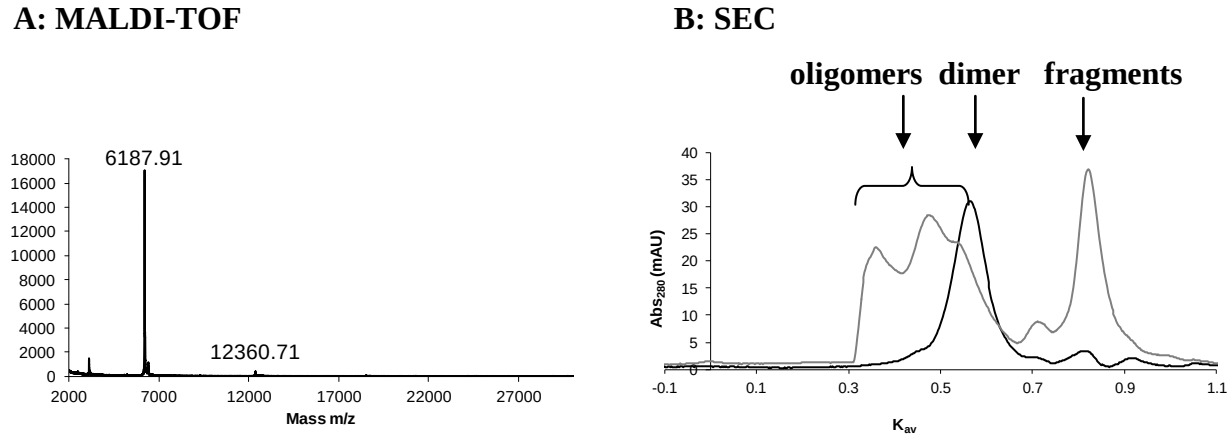


Figure 3.11: Molecular weight analysis of Nterm-Lsr2 post proteolysis. (A) MALDI-TOF MS with major peaks at 6188 Da and 12361 Da which represent the truncated monomer and dimer respectively. (B) oligomerised Nterm-Lsr2 at pH 7.4, 1 mg/mL. Black: full length dimer (19 kDa); grey: Nterm-Lsr2 after trypsin proteolysis (72 kDa, 36 kDa and 24 kDa). Treatment with trypsin removes the N-terminal tag and 3 additional residues, which facilitates the formation of an inter-dimer β-strand. [6, 7]

Nterm-Lsr2 was successfully oligomerised by reproducing the method described in Summers *et al*. [6] The sizes of oligomers formed were smaller than those seen by Summers, which is consistent with the lower protein concentration used in this study. The self-assembly of Nterm-Lsr2 in the crystal structure is linear with each dimeric unit having two identical oligomerisation interfaces at each end (fig. 3.1). With no known capping mechanism to halt assembly, oligomerisation of Nterm-Lsr2 will continue up to the point of protein saturation. [27-29]

3.3.3 Effects of concentration on self-assembly of Nterm-Lsr2

As predicted, increasing the concentration of Nterm-Lsr2 in solution led to an increase in the size of the largest oligomer (fig. 3.12). At pH 7.4 there were three oligomer peaks representing HMW oligomers ranging from a decamer to a 30-mer, medium molecular weight (MMW) oligomers ranging from a hexamer to octamer, and low molecular weight (LMW) oligomers which were always a tetramer. Interestingly, the post digest dimer was never seen on SEC, possibly due to the concentration of the dimer peak being below the

instrument's limit of detection. Increasing, the concentration of Nterm-Lsr2 in solution led to an increase in the size of the largest oligomer (fig. 3.12). At a protein concentration of 4.3 mg/mL, the highest molecular weight oligomer was three times larger than the highest molecular weight oligomer at 1 mg/mL (30-mer versus decamer) and twice as large as the highest molecular weight oligomer at 1.8 mg/mL (30-mer versus tetradecamer).

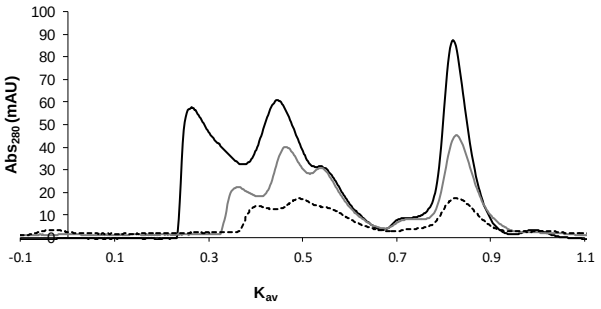


Figure 3.12: Oligomerisation of Nterm-lsr2 at pH 7.4, varying concentrations. Black: 4.3 mg/mL; grey: 1.8 mg/mL and dashed: 1.0 mg/mL. Increasing the concentration leads to the formation of larger oligomers showing that Nterm-Lsr2 assembly is concentration dependent.

As the MMW peak was always the most populated at all concentrations it was believed that the tetramer and MMW oligomers were in a fast equilibrium and the MMW oligomer needed to reach optimal concentration before the HMW oligomers could form.

3.3.4 Equilibrium and distribution of self-assembled supramolecular structures of Nterm-Lsr2

To try and understand the dynamics of protein assembly a number of equilibrium experiments were conducted. It was useful to think of each individual oligomer as a tecton that, once an optimal concentration was reached, could associate into the higher molecular weight species.

Partial digestion with trypsin was used to trigger Nterm-Lsr2 assembly on two separate samples of Nterm-Lsr2 with different concentrations (1.8 mg/mL and 4.3 mg/mL). The

oligomers were separated by size using SEC. Pooled fractions were concentrated to 0.5 mL and re-injected onto the SEC column.

Upon re-injection, the HMW oligomer (30-mer and tetradecamer at 4.3 mg/mL and 1.8 mg/mL respectively) re-equilibrated to a MMW oligomer (octamer and heptamer at 4.3 mg/mL and 1.8 mg/mL respectively) (Appendix 3.5). The apparent uneven number of units was believed to be a result of unwanted trypsin cleavage along the peptide chain and the inclusion of trypsin digest products in the protein assembly. These results agreed with the hypothesis that the HMW and MMW oligomers were in equilibrium whereby the MMW weight oligomer, or tecton, must reach an optimal concentration before higher ordered structures could form. The oligomer concentration dropped due to dilution on the 24 mL column and therefore, upon re-injection, the HMW oligomers dissociated into the MMW species.

The same re-equilibration of oligomers to a LMW species was observed when the MMW oligomers (octamer and heptamer at 4.3 mg/mL and 1.8 mg/mL respectively) were re-injected (Appendix 3.6). The re-injected heptamer (1.8 mg/mL) equilibrated to one species at 30 kDa, approximately the molecular weight of a pentamer. The re-injected octamer (4.3 mg/mL) eluted as three distinct oligomers; an octamer, tetramer and a pentamer.

Unlike the re-injection of the HMW and MMW oligomers described above, the LMW oligomers (tetramer) from the 1.8 mg/mL and 4.3 mg/mL oligomerisation stayed at the same molecular weight when re-injected (Appendix 3.7). The tetramer was also the smallest oligomer seen after proteolysis, suggesting that this was the starting "building block" from which the other oligomers were formed. Results suggest that the HMW species associate from individual tetramers once an optimal or threshold concentration is reached.

When observing the protein assembly at varying concentrations it was noted that the ratio of MMW oligomers to LMW and HMW oligomers stayed static. The ratio was calculated as a percentage of the peak area over the total area of oligomers. For all three concentrations tested (1 mg/mL, 1.8 mg/mL, and 4.3 mg/mL, Appendix 3.5 and 3.6 respectively) the percentage of MMW oligomers was 44 ± 0.07 % (Table 3.2). In contrast, the percentage peak area for the LMW and HMW altered with varying initial protein concentration.

	LMW		MMW		HMW	
Conc	**Peak Area**	**Oligomer size**	**Peak Area**	**Oligomer size**	**Peak Area**	**Oligomer size**
(mg/mL)	**(% volume)**		**(% volume)**		**(% volume)**	
1 .0	29.25	tetramer	43.89	hexamer	26.86	decamer
1.8	34.94	tetramer	44.06	heptamer	21.00	tetradecamer
4.3	16.65	tetramer	43.81	octamer	39.54	30-mer

Table 3.2: Oligomer size and % peak area of the LMW, MMW, and HMW assemblies at varying concentrations. The ratio of oligomers is concentration dependent with the amount of HMW oligomer increasing with decreasing LMW peak area and vice versa. Interestingly, the MMW peak area is approximately the same under all concentrations suggesting that MMW to HMW is the rate limiting step in Nterm-Lsr2 assembly. The tetramer is always the smallest oligomer. The apparent uneven numbers of units are believed to be due to unwanted trypsin cleavage along the peptide chain and the resulting inclusion of trypsin digest products in the protein assembly.

By combining the information from the equilibrium experiments and the integrated peak areas of the different oligomers at varying concentrations, a possible route towards assembly was proposed (fig. 3.13). This is consistent with the MMW oligomer acting as a stable intermediate that reaches an equilibrium population during assembly of the HMW form.

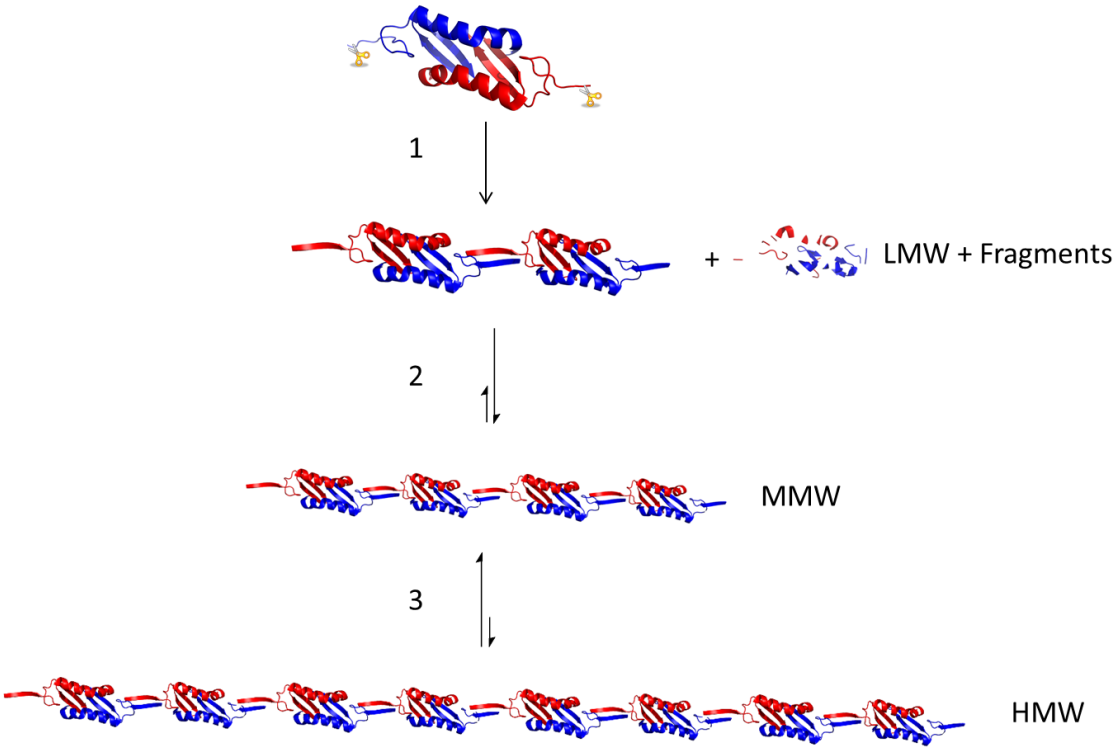


Figure 3.13: Schematic of the suggested equilibrium between the LMW, MMW and HMW oligomers. The tetramer is the smallest oligomer seen post-proteolysis (step 1) and is the starting "building block" from which the other oligomers are formed. Under all concentrations the most abundant oligomer observed is the MMW suggesting a fast equilibrium between LMW and MMW (step 2). The HMW oligomer appears to be formed in a slower equilibrium from the MMW oligomer (step 3).

3.3.5 Effects of pH on self-assembly of Nterm-Lsr2

Variations in pH values have been shown to alter the structure, solubility, activity, and stability of proteins. This is due to the variations in charge of ionisable amino acids above or below their pK_a values. [30] The charges can either contribute to stability, electrostatic interactions or aid in dissociation, electrostatic repulsion. With the additional polar contacts that stabilise the Nterm-Lsr2 dimer interface it was of interest to assess the influence of varying pH conditions on protein oligomerisation. This not only gave insight into the types of interactions taking place in solution, but could also be used as an additional route towards manipulating assembly.

The oligomerisation was conducted at pH 6.0, 8.0, and 9.0 to study how varying this external factor affected the protein assembly. Proteolysis was carried out under neutral pH conditions to ensure consistent enzyme activity. After 10 minutes the pH was adjusted to either pH 6.0, 8.0, or 9.0 by adding 0.5 mL of the appropriate buffer and measuring the pH prior to injecting the sample onto a SEC column. There were small changes in the extent of oligomerisation when varying the pH to 6 or 8, with the largest oligomer being in the MMW range (Table 3.3, Appendix 3.8). However, when the pH was adjusted to 9.0 there was a marked effect on the different types of oligomers seen with only a tetramer forming (fig. 3.14).

pH	LWM	MMW	HMW
6.0	tetramer	hexamer	non
7.4	tetramer	hexamer	decamer
8.0	tetramer	hexamer	non
9.0	tetramer	non	non

Table 3.3: Sizes of oligomers at varying pH. pH 7.4 is the optimum pH for forming large oligomers.

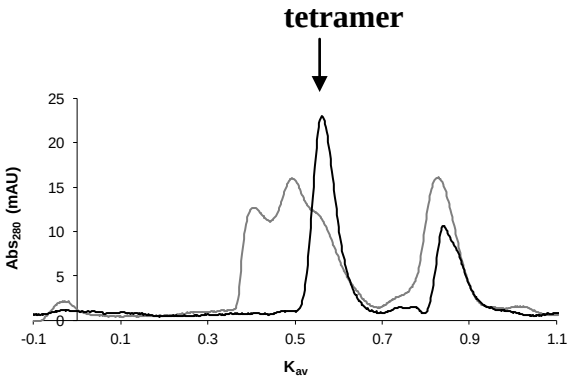


Figure 3.14: The influence of pH on the oligomerisation of Nterm-Lsr2. Grey: Nterm-Lsr2 post-proteolytic cleavage at pH 7.4, 1 mg/mL. Seen is a mixture of HMW assemblies and protein fractions. Black: Nterm-Lsr2 post-proteolytic cleavage at pH 9.0, 1 mg/mL. The only visible oligomer is a tetramer. The assembly is halted by an increase in pH.

Characterisation of the oligomerised protein at pH 9.0 confirmed that the secondary structure was intact (fig. 3.15). Any variations were attributed to the change in secondary structure accompanied by the higher oligomeric state of the protein at pH 9.0, from a dimer to a tetramer.

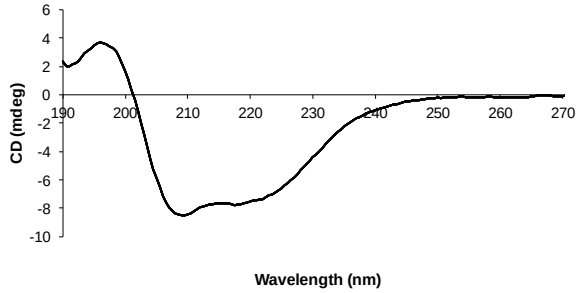


Figure 3.15: CD spectra of Nterm-Lsr2 at pH 9. Troughs at 219 and 209 nm represent α-helix and β-sheet secondary structural elements respectively. Any differences between the pH 9 and pH 7.4 CD spectra (Nterm-Lsr2 pre-proteolysis, fig. 3.10) are the result of changes in the secondary structure accompanied by the higher oligomeric state of the protein at pH 9 (specifically, the loss of random coil and the increase in the number of β-strands).

This pH switching process was also shown to be reversible. The tetramer from the oligomerisation at pH 9.0 was collected and transferred back into a pH 7.4 sodium phosphate buffer using diafiltration. When this was re-injected onto the SEC the tetramer re-equilibrated back into larger oligomers in the range of 24 to 130 kDa (fig. 3.16). To ensure that the effect was not buffer dependent, the samples were analysed in Tris buffer at pH 7.4 and the same pattern of oligomerisation was seen as that with sodium phosphate (Appendix 3.9).

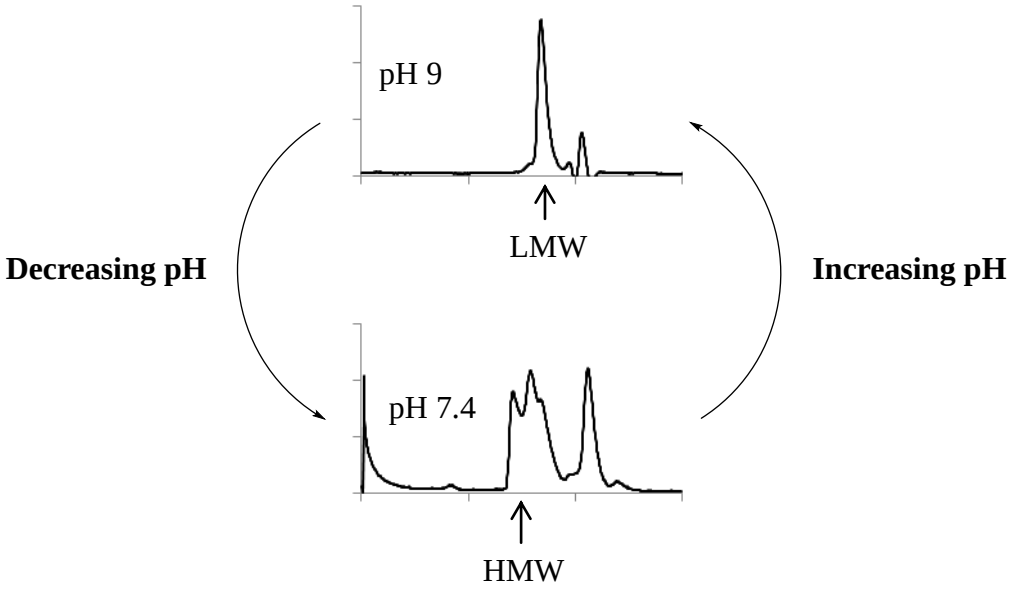


Figure 3.16: Switching between pH 7.4 and pH 9.0 is a reversible process. HMW oligomers are transferred from pH 7.4 buffer into pH 9.0 leading to disassociation into tetramers. This cycle can be a repeated a number of times before the protein becomes unstable.

The stabilisation of the Nterm-Lsr2 oligomers is not only due to the hydrophobic interactions and H-bonding along the β-strand, but also due to the polar contacts from the N-terminal amine (Lys^4). [6] An increase in pH above the pK_a of the residue's alpha-amine (pK_a 8.95) would deprotonate the positively charged amine, thereby removing the polar contacts and significantly destabilising the inter-dimer interface.

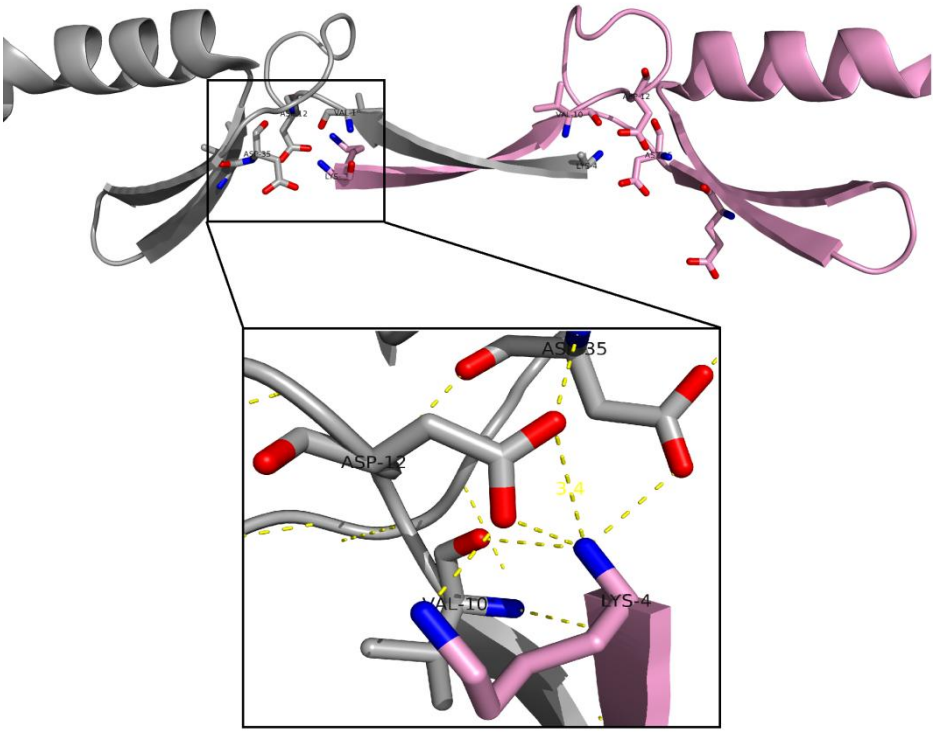


Figure 3.17: Crystal structure demonstrating the inter-dimer interactions. Yellow dotted lines show the electrostatic interactions of the new N-terminal amine and the acidic side chains of Asp^{12} and Asp^{35}, and the carbonyl group of Val^{10} on the neighbouring dimer. Dimer 1 in grey and dimer 2 in pink (one monomer from each dimer is shown for clarity). Deprotonation of the N-terminal amine at pH 9.0 leads to a loss of these interactions, reducing the oligomer stability. *Images generated using Pymol with coordinates from PDB: 4E1R.*

The new interactions that are present at the Nterm-Lsr2 oligomerisation interface (specifically hydrogen bonding from the backbone of the anti-parallel β sheet, hydrophobic interactions between the hydrophobic residues at the dimer-dimer interface, and polar interactions between Lys^4 amine and Asp^{12}, Asp^{35} and Val^{10} on the neighbouring dimer) will all contribute to stabilisation of the oligomer. [31] At pH 9.0 the protein interface becomes less stable due to a loss of the polar contacts. The binding efficiency will decrease with the decreasing number of polar contacts available. Furthermore, there will be a loss in enthalpy gain as the degree of intermolecular bonding decreases, making the formation HMW and MMW oligomers both thermodynamically and kinetically unfavourable.

It is important to note that the tetramer is still present without these electrostatic interactions. In fact, it is the smallest complete protein unit seen, post proteolysis, under all conditions. It is possible that, upon removal of steric hindrance caused by Met^{1}, Ala^{2} and Lys^{3}, the tetramer, rather than dimer, is the most thermodynamically stable oligomer. The equilibrium between dimer, tetramer, MMW and HMW will be sensitive to both protein concentration and the stability of the oligomer interface. The oligomer interface has been destabilised by increasing the pH to 9.0, and therefore a higher concentration will be needed to drive the equilibrium towards forming MMW and HMW oligomers. Thus, only the tetramer is observed.

3.3.6 Discussion

Nterm-Lsr2 undergoes oligomerisation in response to N-terminal cleavage by trace amounts of trypsin. This makes it an attractive target for use as a protein tecton as it has an "off-on" switch at the inter-unit interface and therefore assembly can be initiated on demand. The protocols developed by Summers *et al*. were successfully employed in this study to oligomerise Nterm-Lsr2 upon enzymatic cleavage at Lys^{4}.

By studying the new inter-dimer interface it was possible to predict the conditions needed to influence the assembly. The crystal structures of both Nterm-Lsr2 digested with contaminating proteases (PDB: 4E1P, monoclinic), and Nterm-Lsr2 digested with trypsin (PDB: 4E1R, hexagonal), showed dimers that were joined together at identical interfaces arrange into a linear chain. [6] This was despite the fact that the two crystals were packed into different space groups, which makes it unlikely that the linear arrangement was an artefact of crystallisation. Due to the symmetrical nature of the interfaces, the degree of assembly could be manipulated by varying the initial protein concentration (a 3-fold increase in size from 1 mg/mL to 4.3 mg/mL). The response to protein concentration also allowed the analysis of the assembly equilibrium, with the tetramer and MMW oligomer forming rapidly and the HMW oligomer forming after MMW had reached a critical concentration. The polar contacts at the new interface were also utilised to control assembly, with the protein switching from HMW oligomers to a tetramer in response to increasing the pH above the pK_a of the Lys^{4} amine. Throughout the study it was noted that optimal oligomerisation was being hindered through unwanted trypsin digest along the polypeptide chain. This led to fragmentation of Nterm-Lsr2 which complicated both the analysis and the assembly.

3.4 Ent-Nterm oligomerisation

To improve the triggered assembly of Nterm-Lsr2, a new and unique cleavage site was designed in which the Lys^4 would still be the N-terminal amino acid post-proteolysis, but without any additional fractionation (fig. 3.18, Table 3.4).

Construct Name	Specific cleavage site + intermolecular β-sheet peptide sequence	Total no. Amino acids
Nterm-Lsr2	MSYYHHHHHHDYDIPTTENLYFQGAMAK–KVTVTLV---residues 36 – 86	86
Met-Nterm	M – KVTVTLV---residues 9 – 59	59
Ent-Nterm	MHHHHHHDYKDDDDK – KVTVTLV---residues 23 – 73	73

Table 3.4: Peptide sequences of new cleavage sites at Lys^4.

3.4.1 Designing a specific cleavage site

Initially, it was thought that methionine amino peptidase could be used to cleave at Lys^4 if Ala^2, Lys^3, the TEV linker and the his-tag were removed from the sequence, leaving methionine as the N-terminal amino acid (Met-Nterm). However, it was previously reported that the peptidase activity is vastly reduced when the penultimate amino acid is lysine. [32] Another method of cleavage at Lys^4 was therefore designed to utilise the his-tag linker cleavage site. The sequence was re-engineered so that the N-terminal amino acids Met^1, Ala^2, Lys^3 and the TEV linker were replaced with an enteropeptidase linker. Enteropeptidase linkers have been used in previous studies to attach his-tags to recombinant proteins. [33] The main benefit of using this linker (instead of a TEV cleavage site) is that the entire linker is removed upon proteolysis with enteropeptidase [34] with no extraneous amino acid residues left on the protein.

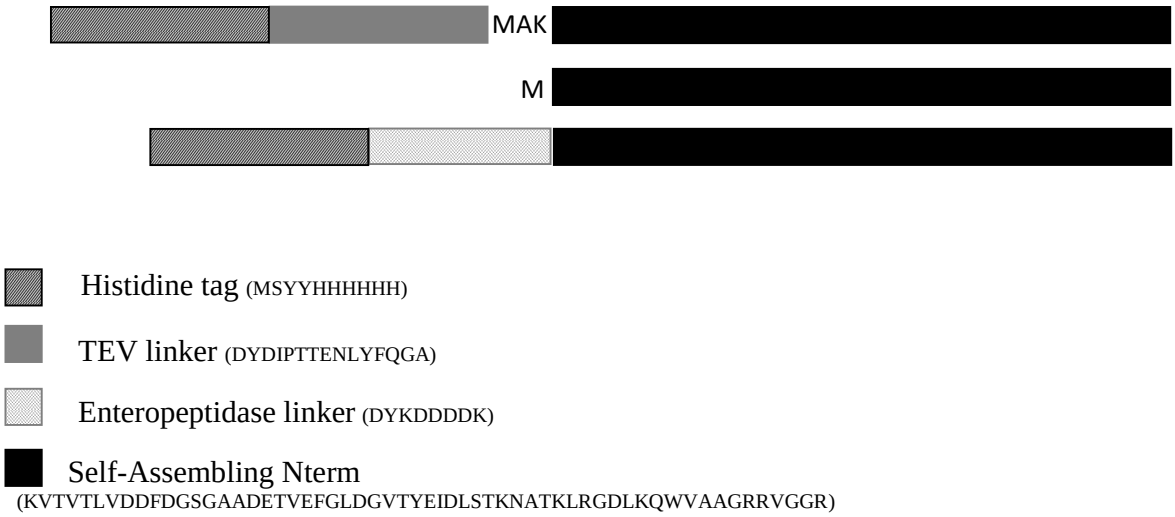


Figure 3.18: Graphical representation of the engineered new cleavage sites at Lys[4]. A new enzyme recognition site has replaced MAK to remove the need for trypsin. This means that there will be no unwanted trypsin cleavage along the polypeptide chain.

3.4.2 Protein expression of Ent-Nterm

Ent-Nterm was expressed in *E. coli* BL21(DE3) cells using standard procedures as described in the methodology (Chapter Two at 2.3.2), and was purified using the same methods employed for Lsr2 and Nterm-Lsr2 purification (fig. 3.19 A). The isolated protein was characterised using SDS-PAGE (fig. 3.19 C).

MALDI-TOF MS showed that the full length protein monomer was 8148.71 Da (fig. 3.20 B), in close alignment to the calculated MW of 8148 Da. Small amounts of dimer could also be seen at 16288 Da. CD analysis showed that the protein was correctly folded with a trough at 209 and 219 nm representing the α-helix and β-sheet respectively (fig. 3.20 A).

Intriguingly, the undigested Ent-Nterm in solution was a tetramer (fig. 3.19 B). This may be due to the reduction of steric constraints between dimers in the absence of Met^1, Ala^2 and Lys^3 which would allow a β-sheet to form. This agrees with the hypothesis that the tetramer is the most stable oligomer when these residues are removed (as was seen with the partially digested Nterm-Lsr2 at pH 9.0).

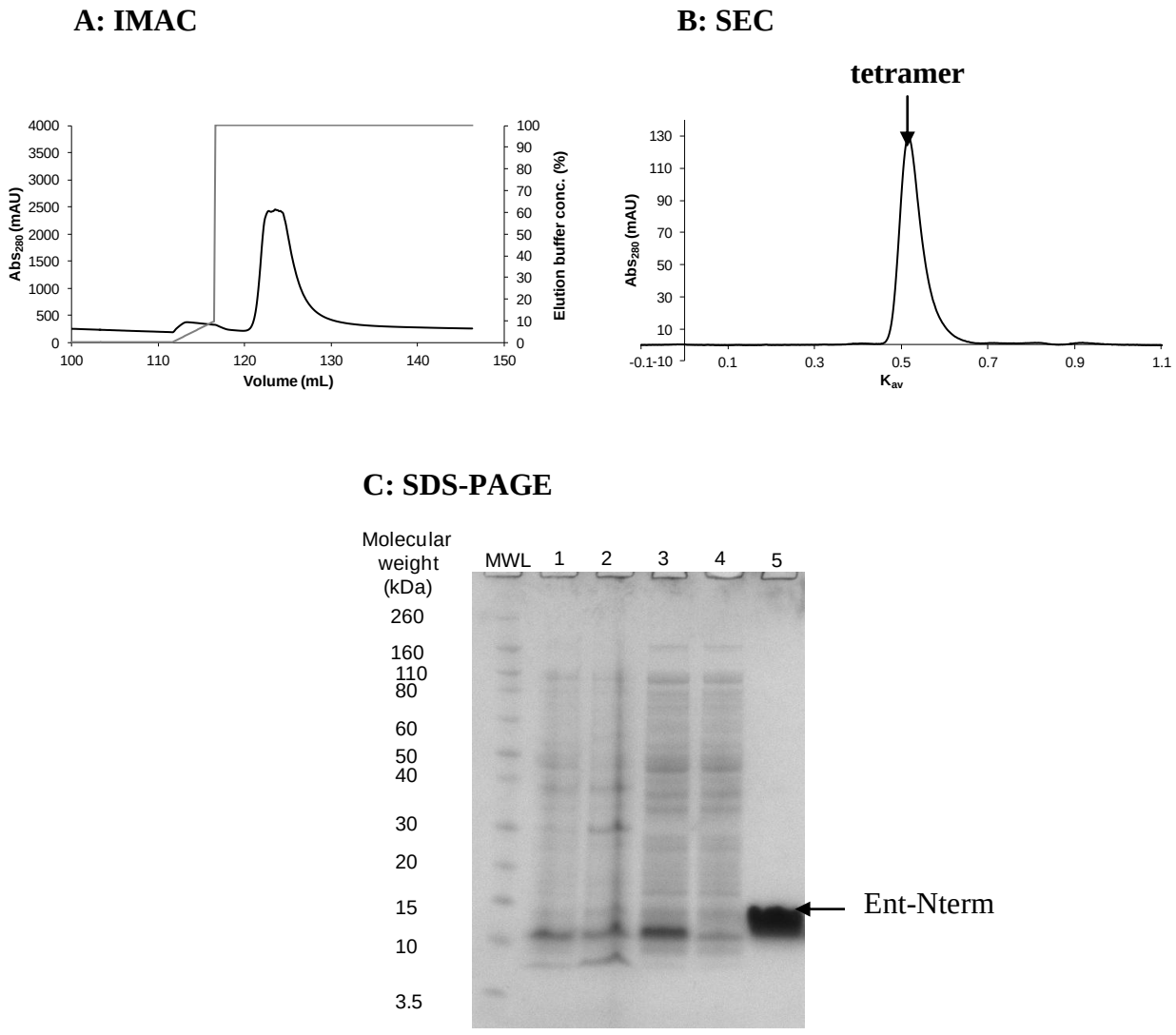


Figure 3.19: Purification and isolation of Ent-Nterm. (A) Elution of Ent-Nterm from a Ni^{2+} IMAC column upon increasing imidazole concentration. Black: absorbance at 280 nm (mAU) and grey: percentage concentration of 1 M imidazole buffer (%). (B) SEC of Ent-Nterm with a peak at 32 kDa, the MW of a tetramer. This is possibly due to absence of Lys3 in this construct, allowing for a single inter-dimer β-strand. (C) SDS-PAGE of protein extraction and purification. 1: crude lysate; 2: insoluble crude; 3: soluble crude, 4: IMAC flow through, 5: isolated Ent-Nterm from IMAC.

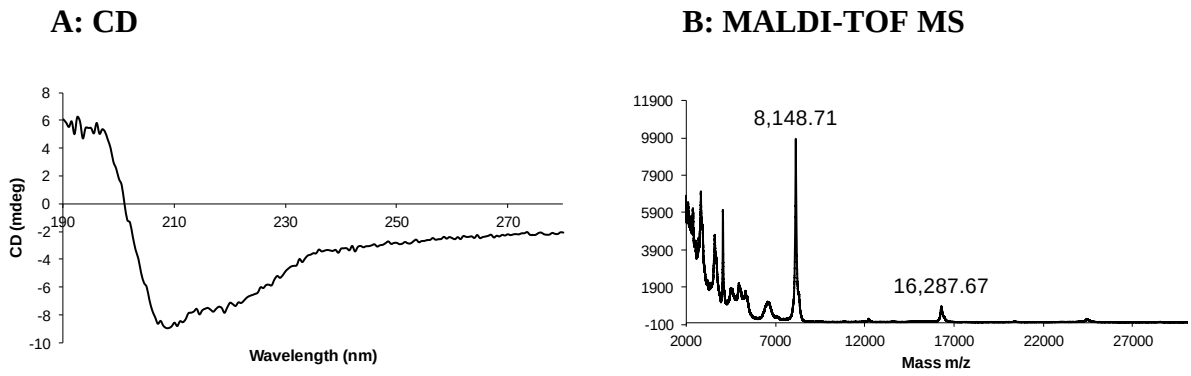


Figure 3.20: Characterisation of Ent-Nterm. (A) CD of full length Ent-Nterm. The troughs at 209 and 219 nm show the secondary structural elements of α-helix and β-sheet respectively. The CD looks similar to that seen for the tetramer at pH 9.0. (B) MALDI-TOF MS of full length Ent-Nterm pre-proteolytic cleavage with enteropeptidase. The major peak is at 8148.71 Da, equal to the calculated MW of an Ent-Nterm monomer (8148 kDa) with a minor peak at the MW of a dimer.

3.4.3 Optimising proteolysis-triggered oligomerisation of Ent-Nterm

Initial experiments were conducted using the ideal conditions for a typical enteropeptidase digest as suggested by the product protocol (1 U/50 µg porcine enteropeptidase /protein, 16h at 22 ^{o}C, Applied Biological Materials (ABM) Inc). Preliminary SEC analysis showed that under these conditions there were HMW oligomers within the sample, but there was also a large amount of undigested protein present (fig. 3.21 A). The concentration of enteropeptidase was increased to 2 or 3 U of enzyme/50 µg of protein. SDS-PAGE and SEC analysis showed that at 3 U/50 µg, a negligible amount of undigested Ent-Nterm was present (fig 3.21 B) and larger oligomers had formed.

The most notable difference between Ent-Nterm and Nterm-Lsr2 was that a much lower initial protein concentration was needed to reach the top end of the HMW oligomer size (28-mer from a 1 mg/mL Ent-Nterm sample). Furthermore, a simpler distribution of assemblies was observed. One at the HMW range (including a 28-mer) and one within the MMW range (octamer). There was also an additional peak that eluted below the lower molecular weight range of the column (10 kDa). MALDI-TOF MS showed that this peak was the monomer post proteolysis (6.2 kDa). It is possible that the monomer was also present after partial

trypsin digest on Nterm-Lsr2 but was masked by the large protein fraction peak. Removal of the N-terminal residues could also cause some disruption along the dimerisation interface.

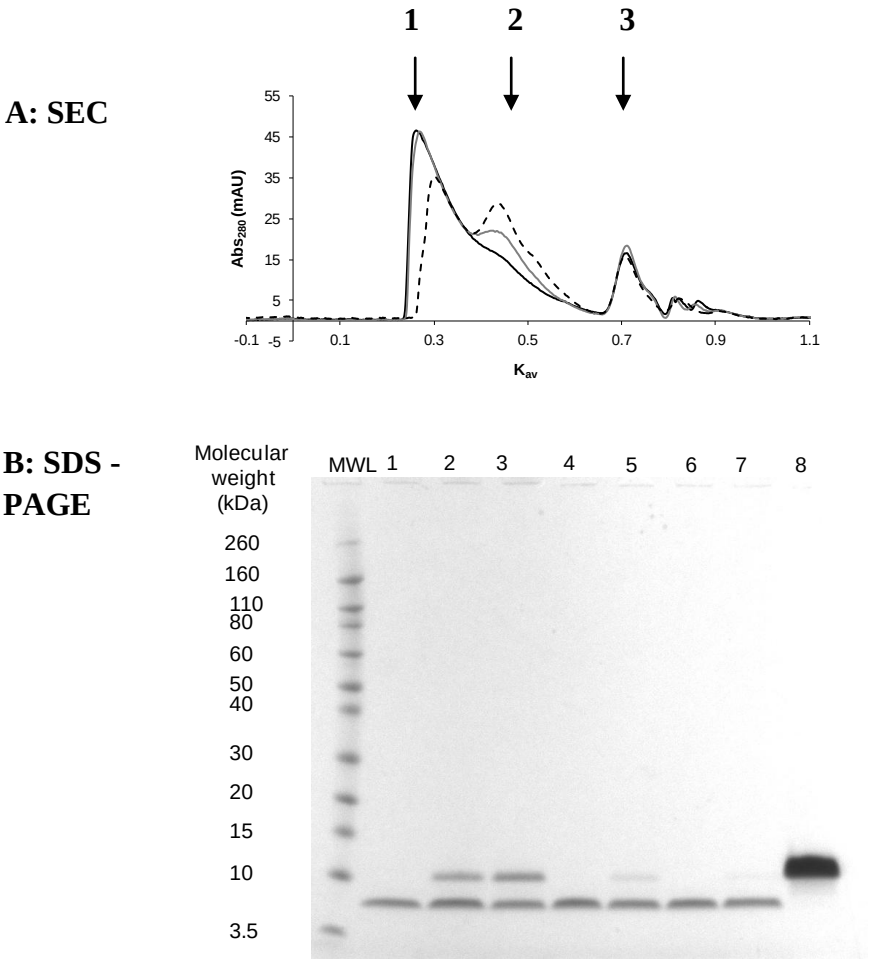


Figure 3.21: Enteropeptidase proteolysis; enzyme concentration trials. (A) SEC of Ent-Nterm post-proteolytic cleavage with varying concentrations of enteropeptidase. Peaks 1, 2, 3 and 4 are HMW, MMW, LMW and monomer respectively. Black: 3 U enteropeptidase/ 50 µg/mL Ent-Nterm; grey: 2 U enteropeptidase/ 50 µg/mL Ent-Nterm; dashed: 1 U enteropeptidase/ 50 µg/mL Ent-Nterm. (B) SDS-PAGE of SEC fractions; 1, 2 and 3: HMW, MMW and LMW respectively of 1 U enteropeptidase/ 50 µg/mL Ent-Nterm proteolysis; 4 and 5: HMW and MMW respectively of 2 U enteropeptidase/ 50 µg/mL Ent-Nterm proteolysis; 6 and 7: HMW and MMW respectively of 3 U enteropeptidase/ 50 µg/mL Ent-Nterm proteolysis and 8: full length Ent-Nterm. Proteolysis with 3 U of enteropeptidase shows negligible amounts of undigested Ent-Nterm and the highest proportion of HMW oligomers.

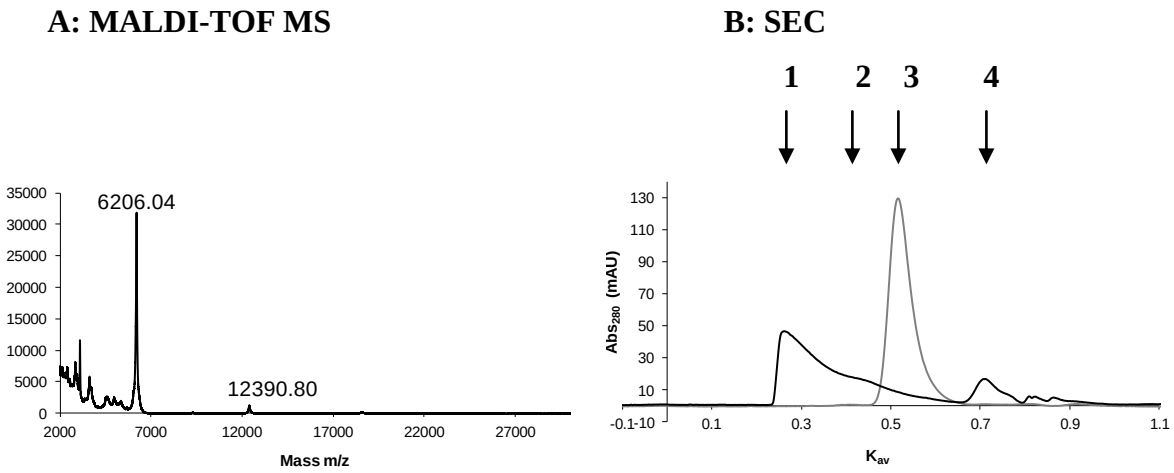


Figure 3.22: Molecular weight analysis of Ent-Nterm post proteolysis. (A) MALDI-TOF MS with major peaks at 6206 Da and 12391 Da which represent the truncated monomer and dimer respectively. (B) Oligomerised Ent-Nterm at 7.4, 1 mg/mL. Grey: full length tetramer (peak 3 - 32 kDa); black: Ent-Nterm after enteropeptidase proteolysis (peak 1 - 174 kDa, peak 2 - 50 kDa and peak 4 – 6 kDa). Treatment with enteropeptidase removes the N-terminal tag which facilitates the formation of continuous inter-dimer β-strands, one on each end of the dimer subunit.

3.4.4 Equilibrium and distribution of self-assembled supramolecular structures of Ent-Nterm

Equilibrium experiments showed that the oligomers behaved in the same way as Nterm-Lsr2 oligomers. The HMW oligomer (28-mer) re-equilibrated to MMW (octamer) species when re-injected onto the column (fig. 3.23 A). Similarly, the MMW oligomer (octamer) re-equilibrated to a slightly lower molecular weight oligomer, which was probably a dilution effect, but the oligomer was still in the MMW range (hexamer) (fig. 3.23 B).

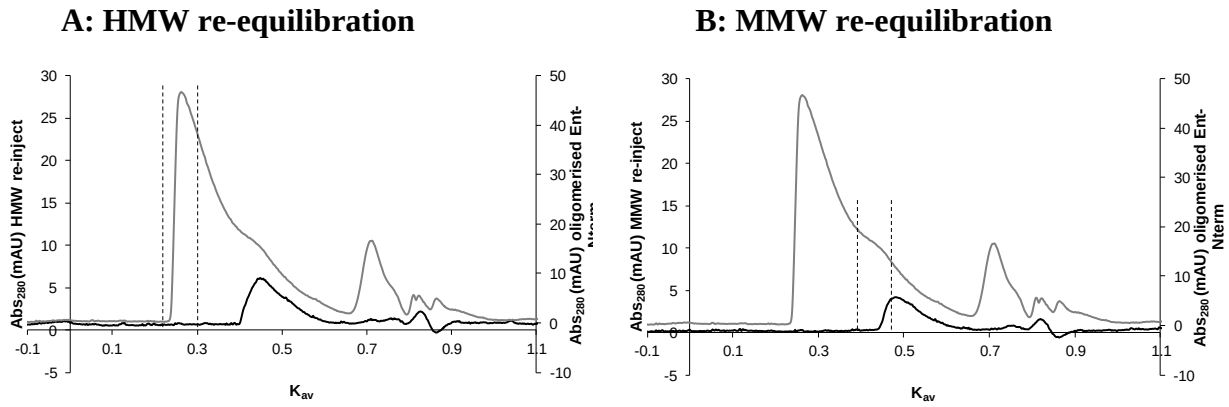


Figure 3.23: Ent-Nterm equilibrium. The eluted peak has been collected, concentrated and re-injected onto the column. (A) HMW peak (28-mer) re-equilibrates to a MMW oligomer (octamer). (B) MMW peak re-equilibrates from an octamer to a hexamer. This is similar equilibrium pattern to Nterm-Lsr2 with the larger oligomers reducing in size upon re-injection onto the gel filtration column. Unlike Nterm-Lsr2, the smallest defined peak is a hexamer.

3.4.5 Effects of pH on self-assembly of Ent-Nterm

The oligomer peak was collected and concentrated to 0.5 mL. This was then exchanged into pH 9.0 buffer using diafiltration. When this was re-injected onto the SEC column the oligomers re-equilibrated to a tetramer (fig. 3.24).

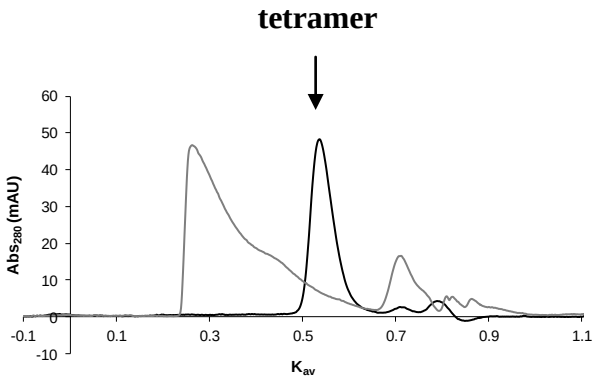


Figure 3.24: The influence of pH on the oligomerisation of Ent-Nterm. Grey: Ent-Nterm post-preoteolytic cleavage at pH 7.4, 1 mg/mL. Seen is a mixture of HMW and MMW assemblies. Black: Ent-Nterm post-preoteolytic cleavage at pH 9.0, 1 mg/mL. Only visible oligomer is a tetramer. Ent-Nterm displays the same behaviour as Nterm-Lsr2 post-proteolysis.

3.4.6 Discussion

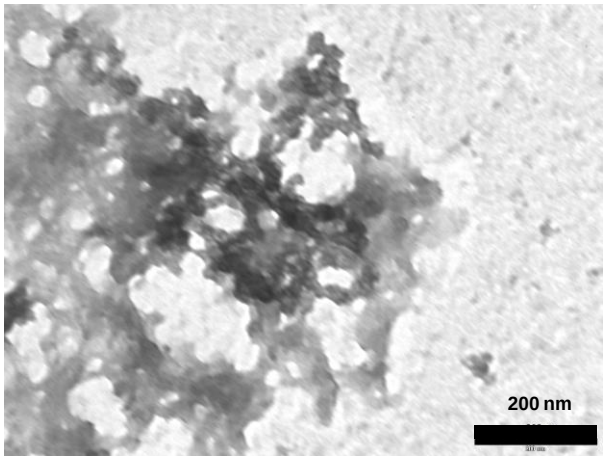
The oligomers formed by enteropeptidase cleavage underwent assembly in the same way as those formed through the partial trypsin digest method, but with a much cleaner distribution of species formed. Decreasing the concentration led to dissociation of the oligomers into smaller species and increasing the pH resulted in a single tetrameric species. It can be inferred, therefore, that the route towards assembly is the same in both systems. The main difference between the two systems was that the large HMW oligomers formed at much lower initial protein concentrations. An explanation for this is that the concentration of oligomerisable protein within the system has increased as there is no longer unwanted proteolysis along the polypeptide chain. As hypothesised, the change in assembly "trigger" led to an improved tecton.

3.5 Transmission electron microscopy

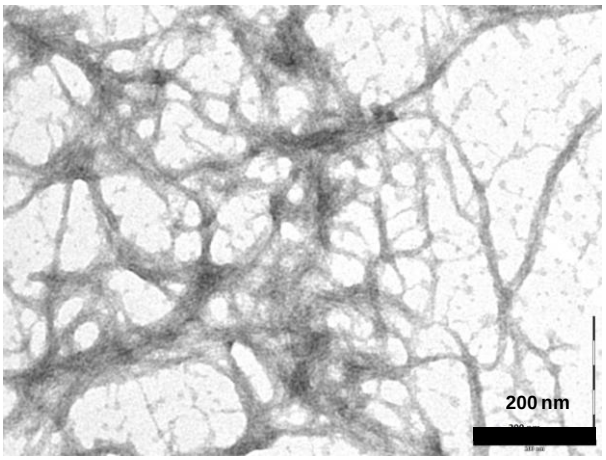
TEM was used to visualize the arrangement of the protein assemblies. The samples were negatively stained and analysed pre and post-proteolysis for both Nterm-Lsr2 and Ent-Nterm. The most striking difference between the samples was the formation of long range "spaghetti-like" structures of Ent-Nterm after treatment with enteropeptidase (fig. 3.25 B). These structures were similar to the networks formed by full length Lsr2 in the presence of DNA. [6] Prior to proteolysis (enteropeptidase treatment for Ent-Nterm and partial trypsin digest for Nterm-Lsr2) the protein existed as disordered aggregate (fig. 3.25 A, C). Upon partial trypsin digest, Nterm-Lsr2 appeared to achieve some long range order. However there was still some protein aggregate present (fig. 3.25 D). This was attributed to the additional fractionation that occurred during the trypsin digest. It is likely that the disordered fragments were entwined with in the ordered network.

The removal of the N-terminal residues at Lys^4 led to the formation of ordered structures rather than a disordered aggregate. It was clear from the TEM images that modifying the his-tag linker cleavage site (and thereby improving the assembly trigger for Nterm-Lsr2) led to the formation of more ordered long range structures without the inclusion of disordered aggregates.

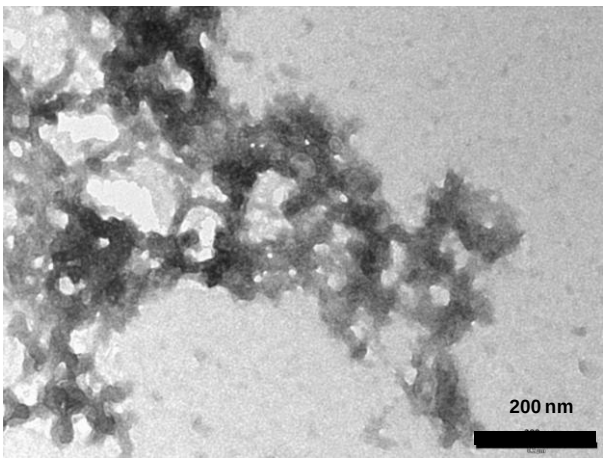
A: Ent-Nterm neat



B: Ent-Nterm post proteolysis



C: Nterm-Lsr2 neat



D: Nterm-Lsr2 post proteolysis

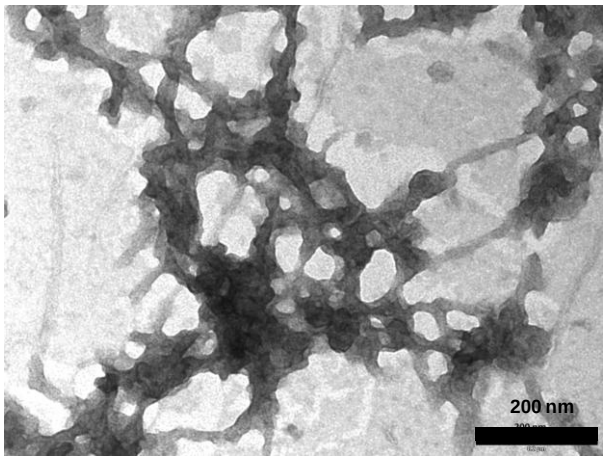


Figure 3.25: TEM of Ent-Nterm and Lsr2-Nterm pre and post-proteolysis. The proteins (50 μg/mL) are negatively stained with uranyl acetate (1%). Black scale bars represent 200 nm. (A) Ent-Nterm pre-proteolysis. Only random aggregate is seen. (B) Ent-Nterm post-proteolysis. An ordered network of strands can be seen. (C) Nterm-Lsr2 pre-proteolysis. Only random aggregate is seen. (D) Nterm-Lsr2 post-proteolysis. There is a degree of order in the network but a lot of aggregate is still visible, likely to be Nterm-Lsr2 fragments. The addition of a specific recognition site at the N-terminus has led to the production of more ordered structures.

3.6 Conclusion

3.6.1 Lsr2

The aim of this research was to develop new methods to utilise proteins for novel assembly systems. Lsr2 was selected as its N-terminal region had been shown to assemble under mild proteolytic conditions. [6] Additional routes towards oligomerisation (namely metal coordination and DNA binding) made Lsr2 an attractive candidate for this project. However, due to the strong DNA binding of the C-terminal region, it was not possible to easily analyse or control assembly. A key feature of an ideal protein tecton is that it should be relatively facile to manipulate and it should assemble in a controllable manner. Without the removal of DNA, Lsr2 is not an ideal protein for forming novel biological nanomaterials.

3.6.2 A novel route for triggered assembly of Nterm-Lsr2

In this study a new route to trigger proteins towards forming higher ordered structures from Nterm-Lsr2 was developed. By engineering the protein to include an enteropeptidase cleavage site, it was possible to avoid fragmentation as a result of trypsin use. The new trigger resulted in an increase in the number of oligomerising units and improved the oligomerisation potential *in vitro*. This was corroborated by TEM, where more ordered, long-range structures were formed from the new tecton.

Variation in protein concentration resulted in different sized oligomers being formed in a stepwise fashion, with the MMW oligomer acting as a stable intermediate between the low and high molecular weight assemblies. pH was used as a switch to toggle between Nterm-Lsr2 tetramers and higher ordered structures. The oligomers formed through enteropeptidase cleavage reacted in the same way to variations in protein concentration and pH as those formed from partial trypsin digest.

With the new enteropeptidase trigger it was possible to form ordered, long-range structures with relative ease. [7] The degree of oligomerisation was regulated through external factors which gave additional control of the system. Downstream, this new tecton trigger could be used as the basis of new responsive materials, biosensors, electronics, or drug delivery vehicles.

3.7 References

1. Bhattacharya P, Du D, Lin YH: Bioinspired nanoscale materials for biomedical and energy applications. *Journal of the Royal Society Interface* 2014, 11(95).
2. Gradisar H, Jerala R: Self-assembled bionanostructures: proteins following the lead of DNA nanostructures. *Journal of Nanobiotechnology* 2014, 12(1):4.
3. Anfinsen CB: Formation and stabilisation of protein structure. *Biochemical Journal* 1972, 128(4):737-&.
4. Alm E, Arkin AP: Biological networks. *Current Opinion in Structural Biology* 2003, 13(2):193-202.
5. Nicolau DE, Phillimore J, Cross R, Nicolau DV: Nanotechnology at the crossroads: the hard or the soft way? *Microelectronics Journal* 2000, 31(7):611-616.
6. Summers EL, Meindl K, Uson I, Mitra AK, Radjainia M, Colangeli R, Alland D, Arcus VL: The structure of the oligomerization domain of Lsr2 from Mycobacterium tuberculosis reveals a mechanism for chromosome organization and protection. *PloS ONE* 2012, 7(6).
7. Ashmead HM, Negron L, Webster K, Arcus V, Gerrard JA: Proteins as supramolecular building blocks: Nterm-Lsr2 as a new protein tecton. *Biopolymers* 2015, 103(5):260-270.
8. Ali SS, Xia B, Liu J, Navarre WW: Silencing of foreign DNA in bacteria. *Current Opinion in Microbiology* 2012, 15(2):175-181.
9. Browning DF, Grainger DC, Busby SJW: Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Current Opinion in Microbiology* 2010, 13(6):773-780.
10. Gordon BRG, Imperial R, Wang L, Navarre WW, Liu J: Lsr2 of Mycobacterium represents a novel class of H-NS-like proteins. *Journal of Bacteriology* 2008, 190(21):7052-7059.
11. Colangeli R, Haq A, Arcus VL, Summers E, Magliozzo RS, McBride A, Mitra AK, Radjainia M, Khajo A, Jacobs WR *et al*: The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proceedings of the National Academy of Sciences of the United States of America* 2009, 106(11):4414-4418.
12. Qu Y, Lim CJ, Whang YR, Liu J, Yan J: Mechanism of DNA organization by Mycobacterium tuberculosis protein Lsr2. *Nucleic Acids Research* 2013, 41(10):5263-5272.
13. Gordon BRG, Li YF, Cote A, Weirauch MT, Ding PF, Hughes TR, Navarre WW, Xia B, Liu J: Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proceedings of the National Academy of Sciences of the United States of America* 2011, 108(26):10690-10695.
14. Gordon BRG, Li YF, Wang LR, Sintsova A, van Bakel H, Tian SH, Navarre WW, Xia B, Liu J: Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America* 2010, 107(11):5154-5159.
15. Chen JM, Ren H, Shaw JE, Wang YJ, Li M, Leung AS, Tran V, Berbenetz NM, Kocincova D, Yip CM *et al*: Lsr2 of Mycobacterium tuberculosis is a DNA-bridging protein. *Nucleic Acids Research* 2008, 36(7):2123-2135.
16. Fágáin CÓ: Understanding and increasing protein stability. *Biochimica et Biophysica Acta* 1995, 1252(1):1-14.

17. Arakawa T, Prestrelski SJ, Kenney WC, Carpenter JF: Factors affecting short-term and long-term stabilities of proteins. *Advanced Drug Delivery Reviews* 1993, 10(1):1-28.
18. Fields PA: Review: Protein function at thermal extremes: balancing stability and flexibility. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* 2001, 129(2–3):417-431.
19. Foo CWP, Bini E, Hensman J, Knight DP, Lewis RV, Kaplan DL: Role of pH and charge on silk protein assembly in insects and spiders. *Applied Physics A* 2006, 82(2):223-233.
20. Sapsford KE, Pons T, Medintz IL, Higashiya S, Brunel FM, Dawson PE, Mattoussi H: Kinetics of metal-affinity driven self-assembly between proteins or peptides and CdSe-ZnS quantum dots. *Journal of Physical Chemistry C* 2007, 111(31):11528-11538.
21. Brodin JD, Carr JR, Sontz PA, Tezcan FA: Exceptionally stable, redox-active supramolecular protein assemblies with emergent properties. *Proceedings of the National Academy of Sciences of the United States of America* 2014, 111(8):2897-2902.
22. Bai YS, Luo Q, Zhang W, Miao L, Xu JY, Li HB, Liu JQ: Highly ordered protein nanorings designed by accurate control of Glutathione S-Transferase self-assembly. *Journal of the American Chemical Society* 2013, 135(30):10966-10969.
23. Ardini M, Giansanti F, Di Leandro L, Pitari G, Cimini A, Ottaviano L, Donarelli M, Santucci S, Angelucci F, Ippoliti R: Metal-induced self-assembly of Peroxiredoxin as a tool for sorting ultrasmall gold nanoparticles into one-dimensional clusters. *Nanoscale* 2014, 6(14):8052-8061.
24. Blommel PG, Fox BG: A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expression and Purification* 2007, 55(1):53-68.
25. Ramelot TA, Raman S, Kuzin AP, Xiao R, Ma L-C, Acton TB, Hunt JF, Montelione GT, Baker D, Kennedy MA: Improving NMR protein structure quality by Rosetta refinement: A molecular replacement study. *Proteins-Structure Function and Bioinformatics* 2009, 75(1):147-167.
26. Dupre M, Cantel S, Martinez J, Enjalbal C: Occurrence of C-terminal residue exclusion in peptide fragmentation by ESI and MALDI tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry* 2012, 23(2):330-346.
27. Morrow JS, Marchesi VT: Self-assembly of spectrin oligomers in vitro: a basis for a dynamic cytoskeleton. *The Journal of Cell Biology* 1981, 88(2):463-468.
28. Fung SY, Keyes C, Duhamel J, Chen P: Concentration effect on the aggregation of a self-assembling oligopeptide. *Biophysical Journal* 2003, 85(Biophysical Journal):537-548.
29. Hong Y, Pritzker MD, Legge RL, Chen P: Effect of NaCl and peptide concentration on the self-assembly of an ionic-complementary peptide EAK16-II. *Colloids and Surfaces B: Biointerfaces* 2005, 46(3):152-161.
30. Grimsley GR, Scholtz JM, Pace CN: A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Science* 2009, 18(1):247-251.
31. Valery C, Pandey R, Gerrard JA: Protein beta-interfaces as a generic source of native peptide tectons. *Chemical Communications* 2013, 49(27):2825-2827.
32. Benbassat A, Bauer K, Chang SY, Myambo K, Boosman A, Chang S: Processing of the initiation methionine from proteins - properties of the escherichia-coli methionine aminopeptidase and its gene structure. *Journal of Bacteriology* 1987, 169(2):751-757.
33. Mikhailova AG, Rumsh LD: Enteropeptidase - Structure, function, and application in biotechnology. *Applied Biochemistry and Biotechnology* 2000, 88(1-3):159-174.

34. Maroux S, Baratti J, Desnuell.P: Purification and specificity of porcine enterokinase. *Journal of Biochemistry* 1971, 246(16):5031-&.

Chapter Four: High molecular weight assembly of *Hs*Prx3-6his

4.1 Introduction

Prxs can adopt a wide range of highly ordered quaternary structures, from dimers and toroids [1, 2], to stacks and tubes [3, 4], catenanes [5], and, under certain conditions, cages (Chapter 1 at 1.6.8). [6] Many members of the Prx family have the capacity to function not only as a peroxidase [7, 8] but also a chaperone, [9] a H_2O_2 regulator, [10] and other physiological roles (Chapter 1 at 1.6.7). [11] This ability to function in multiple cellular roles is intrinsically linked to the oligomeric state of the protein. [12] With the wealth of understanding available in the literature on how these different oligomers form, it is possible to predict conditions that can initiate controlled assembly of biological nanoscale materials. [4, 13] For this reason, Prx has been identified as an ideal candidate for forming novel structures in the field of protein nanotechnology. [4]

The first step towards generating nanoscale materials from proteins is designing a stable tecton from which more complex structures can assemble. In the case of human Prx 3 (*Hs*Prx3), this is the LMW dodecameric ring, which is formed from six oblique homodimers. [4] There are a number of ways of stabilising the ring that have been seen in the literature (Chapter 1 at 1.6.6). One such method involves adding a reducing agent which prevents the active site disulfide bond from forming, a key event in the disassociation of the ring. [1, 4, 7] Other methods incorporate site specific mutations, which lock the ring in place. [14] The pH conditions can also influence the toroid stability. A histidine residue at the inter-dimer (A-type interface), conserved in many typical 2-cys Prx, forms an salt bridge (with arginine on the adjacent dimer) that locks the C_p loop in the closed conformation in pH conditions below the histidine side group pK_a (pH 7.37). [15] This is believed to stabilise the LMW structure at pH 7.3 and below. [15] It has been hypothesised that histidine is involved in stabilising the A-type interface of other peroxidredoxins that exist as oxidised toroids above pH 7.3. For example, the crystal structure of *Ancylostoma ceylanicum* Prx 1 (*Ace*Prx-1) showed that His^{107} can form electrostatic interactions with backbone oxygen on the adjacent dimer as well as π-π stacking with His^{107}'. [16] This residue is conserved in human Prx 4 (*Hs*Prx4), which is also a stable oxidised toroid above pH 7.3. [17]

Intriguingly, the addition of an N-terminal affinity tag, often cleaved after protein purification, can have a profound effect on the oligomeric state of Prx (Chapter One at 1.6.6). In the case of histidine-tagged bovine Prx 3 (*Bt*Prx3), which shares a 93% sequence homology with *Hs*Prx3, the dodecamer is stabilised in both the reduced and oxidised form. [18] Additional non-native residues at the N-terminus of the protein have also been shown to have a direct influence on the geometry of the ring. [19] Furthermore, histidine tags have been used to aid the assembly of the LMW tecton into the HMW stacks and rings via divalent metal coordination. [13]

It is clear that histidine in the native protein sequence is involved in the quaternary structure assembly of many Prxs. [15-17] Furthermore, the presence of a histidine tag can aid the controlled assembly of LMW Prx into its HMW form. The evidence that the tags can not only be utilised directly via metal coordination, but also have the capacity to change the ring geometry and quaternary structure of Prx indicates that there is an abundance of potential influences yet to be explored. The following chapter will document how the presence of an N-terminal poly-6-histidine tag on *Hs*Prx3 (*Hs*Prx3-6his) affects the structure and behaviour of the protein, and the novel methods that have been employed to harness these changes to drive assembly into complex biological nanomaterials.

4.2 Expression and purification of *Hs*Prx3-6his

The gene encoding *Hs*Prx3-6his, cloned into pET151-D-TOPO, was synthesised by Epoch Life Science Inc. Chemically competent *E. coli* cells (BL21 (DE3) ROSETTA, Novagen, 2013; genotype – F- *ompT hsdS*B[B- mB-] *gal dcm* [DE3] pRARE [CamR]) were transformed with the plasmid using heat shock (Chapter 2 at 2.2.2.2). These cells were chosen due to the pRARE codon, which supplied tRNA for the rare codons of *Hs*Prx3 and the presence of T7 RNA polymerase, which allowed expression from the pET vector. The protein was successfully expressed in standard lysogeny broth (LB) media (Chapter 2 at 2.3.1.1) after shaking (180 rpm) at 37^{o}C for 4 hours, in the presence of ampicillin and chloramphenicol, and then at 26^{o}C (with the addition of IPTG) for a further 20 hours.

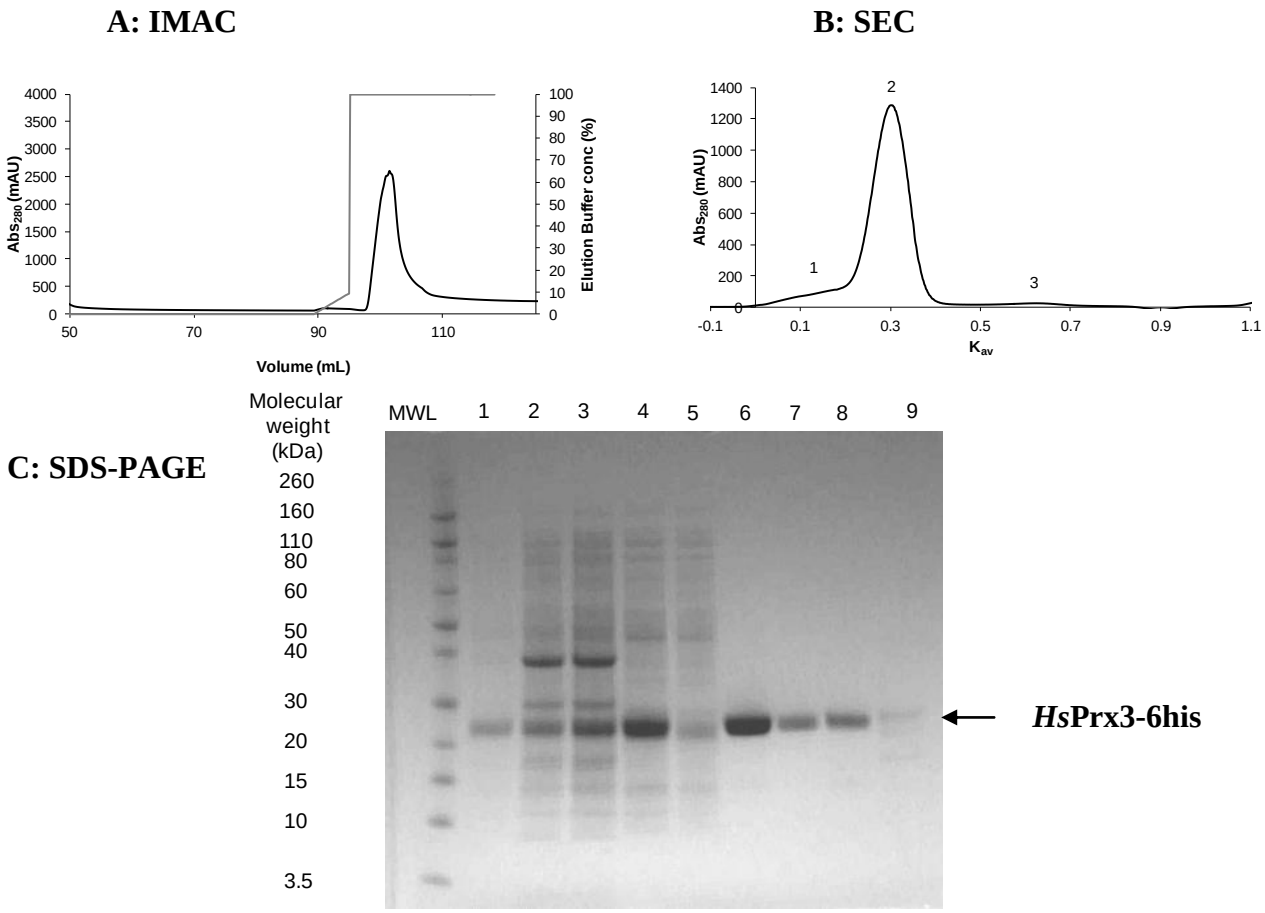


Figure 4.1: Purification of *Hs*Prx3-6his: (A) Elution of *Hs*Prx3-6his from a Ni^{2+} IMAC column upon increasing imidazole concentration. Black: absorbance at 280 nm (mAU); grey: percentage concentration of 0.5 M imidazole buffer (%). (B) SEC chromatography of eluted IMAC peak. The major peak elutes as a dodecamer (peak 2) with a small shoulder (peak 1) and a peak at lower MW containing impurities and small amount of *Hs*Prx3-6his (peak 3). (C) SDS-PAGE gel of *Hs*Prx3-6his purification. 1: crude lysate; 2: insoluble crude 1; 3: insoluble crude 2; 4: soluble crude; 5: IMAC flow-through; 6: IMAC imidazole eluted fraction; 7: SEC peak 2; 8: SEC peak 2; 9: SEC peak 3.

Cells were lysed using sonication (Chapter 2 at 2.3.2.2) and the insoluble fraction was removed by centrifugation (18 000 x g, 30 minutes, 4 ^{o}C). *Hs*Prx3-6his was isolated from the soluble crude broth using IMAC and the sample was separated from any remaining impurities using SEC (fig. 4.1 A and B). SEC data showed a major peak around 305 kDa, close to the calculated molecular weight of the *Hs*Prx3-6his dodecamer (304 kDa), flanked by a HMW shoulder. Purified *Hs*Prx3-6his was in both of these peaks, as identified by a band at 25 kDa

(*Hs*Prx3-6his monomer) on SDS-PAGE (fig. 4.1 C). It is possible that trace amounts of Ni^{2+} had been leached from the IMAC column and coordination to the histidine tag was driving the formation of the HMW species. This phenomenon has been noted in previous studies. [4, 13, 18]

During the final purification step of SEC, the protein was transferred into storage buffer (Chapter 2 at 2.3.2.6) which contained a chelating agent, 10 mM EDTA, and reducing agent, 2 mM TCEP. After incubation at 4^{o}C in the presence of EDTA for at least an hour, and the subsequent removal of the chelating agent, no shoulder peak was observed (fig. 4.2).

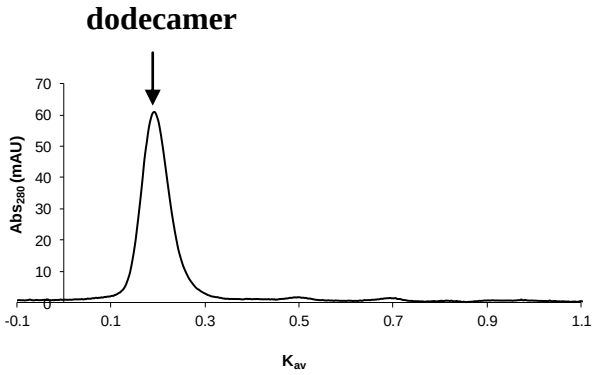


Figure 4.2: SEC of *Hs*Prx3-6his after removal of EDTA and Ni^{2+} in reducing conditions (150 mM NaCl, 20 mM HEPES, 2 mM TCEP, pH 8.0, 1 mg/mL). The shoulder peak is no longer present and the oligomer is in the LMW form (305 kDa, dodecamer) suggesting that the presence of the HMW peak was due to Ni^{2+} coordination.

4.3 N-terminal 6-histidine tag stabilises the toroid in non-reducing conditions

Like many other typical 2-cys Prxs (Chapter 1, at 1.6.6), [2, 5, 20] wild type human Prx 3 (*Hs*Prx3-WT) switches between a dimer and a dodecamer under non-reducing and reducing conditions respectively. [4] Without a reducing agent present, a disulfide bond can form between the peroxidatic cysteine (Cys_p) and the resolving cysteine (Cys_r) on the neighbouring subunits. [2, 7, 21] For this to occur, the C_p loop must undergo local unfolding [22] which, in turn, displaces key residues at the A-type interface causing the toroid to disassemble (Chapter 1 at 1.6.6, fig. 1.15). [2]

While the N-terminus of the protein is not located at the A-type interface, the presence of additional amino acids at this point has been shown to influence the toroid stability (Chapter 1 at 1.6.6). [17-19, 23] With the evidence from the literature, it is apparent that there are routes towards toroid stabilisation that do not involve direct mutations of the A-type interface. Through the size and structure characterisation techniques employed in this work, it has been established that the presence of an N-terminal histidine tag on *Hs*Prx3 generates a stable dodecamer in both reducing and non-reducing conditions. This provides an ideal tecton from which larger structures can be formed.

4.3.1 Analysis of *Hs*Prx3-6his under non-reducing conditions

Size Exclusion Chromatography

*Hs*Prx3-6his was transferred into non-reducing buffer at pH 8.0 (Chapter 2 at 2.3.2.5) which removed the reducing agent, the EDTA, and any trace metals from the protein sample. This was to allow for the formation of the intra-dimer disulfide bond and to eliminate the possibility of oligomerisation via histidine tag to metal ion chelation. The sample was left O/N to allow the protein rings to dissociate in accordance with protocols from the literature. [4] After this time, the protein sample was analysed using a gel filtration column pre-equilibrated with the non-reducing pH 8.0 buffer (Chapter Two at 2.5.1). Unlike the wild type untagged construct, [4] non-reduced *Hs*Prx3-6his eluted at 305 kDa. This shows that, even without the addition of a reducing agent, the dodecamer does not dissociate into its dimeric form (fig. 4.3).

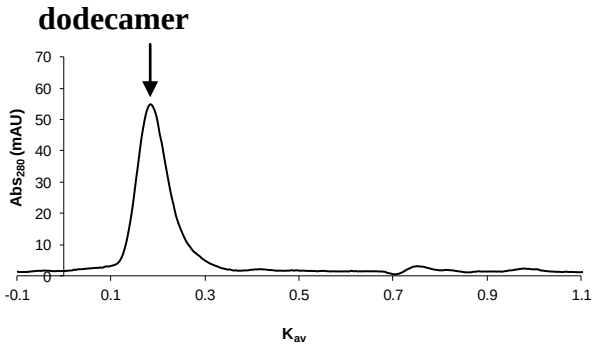


Figure 4.3: SEC of *Hs*Prx3-6his after removal of EDTA and Ni^{2+} in non-reducing conditions (150 mM NaCl, 20 mM HEPES, pH 8.0, 1 mg/mL). Unlike *Hs*Prx3-WT, [4] the protein exists predominantly as a dodecamer.

This is consistent with the findings of Cao *et al*. who found that the *Bt*Prx3-tagged dodecamer is stable under non-reducing conditions at 1 mg/mL, whereas the untagged construct only exists as a mixture of dimer and dodecamer at concentrations in excess of 10 mg/mL. [18] Additional conditions were also trialled (with varied imidazole and EDTA concentration, incubation time, and low protein concentration) but none were successful at destabilising the dodecamer (Appendix. 4.1). The untagged construct (*Hs*Prx3-WT) dissociates under non-reducing conditions, without the addition of chelating agents, at 4 ^{o}C and at 1 mg/mL. [4] The only major difference between the two constructs is the presence of the tag and linker. Therefore, the stability of *Hs*Prx3-6his dodecamer under non-reducing conditions is likely to be a consequence of the additional sequence at the N-terminus of the protein.

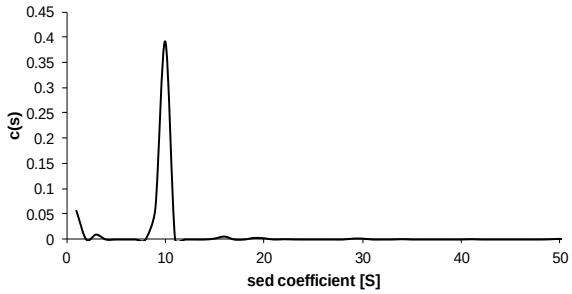
Analytical Ultracentrifugation and Transmission Electron Microscopy

While gel filtration was a good starting point, used to confirm toroid stability, additional analysis was needed to confirm the accurate size and structure of *Hs*Prx3-6his. During gel filtration there is always a degree of dilution due to the diffusion of the sample as it passes through the column. [24] It is possible that there is a low concentration of dimer present that may not be visible on SEC.

Samples were prepared for the AUC in the same way as for the gel filtration column. The concentration was reduced to take gel filtration dilution into account while maintaining an OD greater than 0.1 at 280 nm, necessary for detection in the AUC. [25] As a comparison between the reduced and non-reduced form, the sample was also analysed in the presence of a

reducing agent (2 mM TCEP) which showed a single species at the molecular weight of a dodecamer (appendix 4.2).

A: AUC



B: TEM

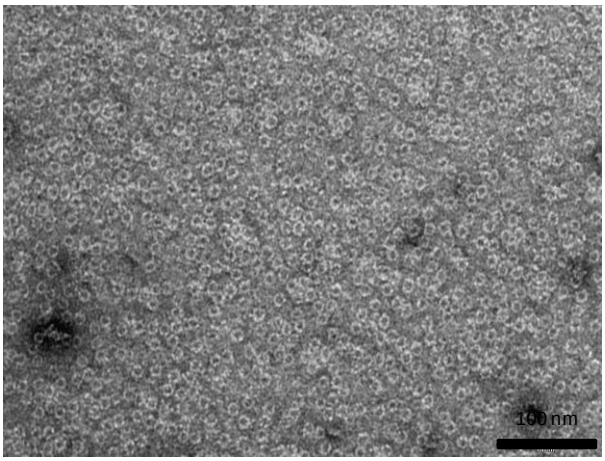


Figure 4.4: Further size and structural analysis of *Hs*Prx3-6his under non-reducing conditions (150 mM NaCl, 20 mM HEPES, pH 8.0). (A) AUC of *Hs*Prx3-6his shows a single species corresponding to the molecular weight of 300 kDa, close to the calculated weight of the dodecamer. (B) TEM of the *Hs*Prx3-6his, diluted to 50 μg/mL in dH_2O prior to grid preparation and stained with 1% uranyl acetate, (140 k x magnification) shows that the protein is in the LMW (dodecameric) form. This further demonstrates that the presence of the histidine tag is stabilising the toroid under non-reducing conditions.

For the non-reduced sample there was a predominant species around 310 kDa (fig. 4.4 A), with a sedimentation coefficient of 11.19 and a frictional ratio of 1.41 (Equation 1). [26] There was also a small peak around the molecular weight of a dimer and a number of peaks

relating to HMW oligomers but all of these were < 2% of the peak volume of the dodecamer peak.

$$M = \frac{s \cdot f \cdot Na}{1 - \rho \cdot \bar{\upsilon}}$$

M = mass
s = sedimentation coefficient
f = frictional coefficient
N_a = Avogadro's number
ρ = solvent density
$\bar{\upsilon}$ = partial specific volume

Equation 4.1

The trace amounts of dimer seen by AUC that were not visible using gel filtration indicate a small concentration dependence for dodecamer stability. However, the predominant species was still a dodecamer with the ratio (~98:2) between the two oligomers. This is far greater than that seen in other studies at the same protein concentration. [2] Therefore the tag significantly improves the stability of the dodecamer under non-reducing conditions.

The structure of the dodecamer was confirmed using TEM (fig. 4.4 B). The sample was diluted to 50 μg/mL in dH_2O directly before grid preparation and stained with uranyl acetate (1 %). It is not possible to observe dimers on TEM due to the resolution being too low to image such a small structure. Unlike previous studies, there were no double stacks seen using TEM, [4] likely due to the additional step in this study of metal chelation after IMAC purification.

4.3.2 The formation and disruption of the active site disulfide bond

Non-reducing SDS-PAGE

A disulfide bond between the active site thiols forms as a reaction intermediate during the catalytic cycle (Chapter 1 at 1.6.1). [1, 21, 27] This is later reduced by a thiol reductant to regenerate the peroxidatic cysteine. [2, 7, 12] Due to the covalent nature of this intra-dimer bond, the oligomer is SDS-resistant and the denatured protein appears as a dimer band (50 kDa) on non-reduced SDS-PAGE electrophoresis. Reducing agents can break the disulfide bond, or site specific mutations of peroxidatic cysteines can prevent the formation of disulfide bonds, resulting in the a 25 kDa protein band as expected for a monomer. [28]. The formation of the disulfide bond is preceded by local unfolding of the active site which, in turn, displaces key residues at the A-type interface resulting in dissociation of the toroid into dimeric units (Chapter 1 at 1.6.6). *Hs*Prx3-6his is a stable toroid under non-reducing conditions, so it is important to ensure that the disulfide bond is still able to form. The protein, which had not been pre-treated with EDTA, was analysed with or without reducing agent (± 2 mM TCEP) and under varying chelating conditions (± 10 mM EDTA, ± 20 mM imidazole). The different chelating conditions were analysed as it was hypothesised that trace amounts of metal could coordinate to the histidine tags. This would further stabilise the dimer, giving a false positive result. Bovine serum albumin (BSA) was used as a positive control to show that protein oligomers can be seen on non-reduced SDS-PAGE (fig. 4.5, lane 1).

A mixture of dimer (~ 50 kDa) and monomer (~ 25 kDa) was seen for *Hs*Prx3-6his in the presence of a reducing agent. The dimer band was weakened with the addition of imidazole and completely disappeared with EDTA and EDTA + imidazole (fig. 4.5 lanes 2 - 5). It is possible that trace amounts of Ni^{2+} in the sample were helping to stabilise the dimer through coordination to the histidine tag. Once the metal ions were removed, and with no disulfide bond at the B-type interface, the oligomer disassembled upon denaturing with heat.

For all the non-reduced samples, only a dimer (~ 50 kDa) was seen on SDS-PAGE, even after heating the protein at 95 ^{o}C for 60 seconds (fig. 4.5, lanes 6 – 9). Unlike the reduced samples, the addition of chelating agents did not disrupt the strong intra-dimer interactions. This showed that the dimer interface was stable, indicating the formation of a disulfide bond as seen in previous work on *Hs*Prx3. [28]

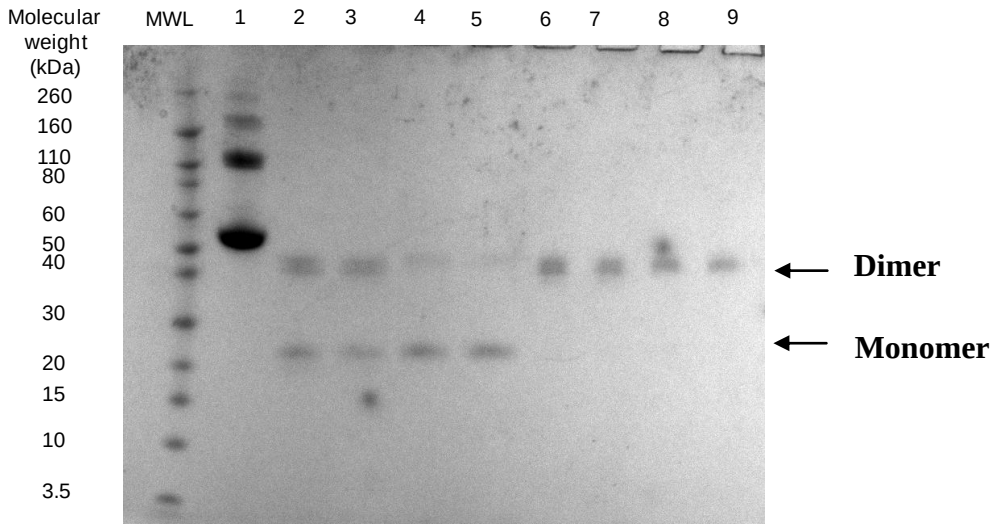


Figure 4.5: SDS-PAGE of *Hs*Prx3-6his under varying conditions. Wells 2 – 5 contain 2 mM TCEP (red) and wells 6 - 9 are non reduced (NR). 1: BSA oligomer positive control; 2: *Hs*prx3-6his red; 3: *Hs*prx3-6his red + 20mM imidazole; 4: *Hs*prx3-6his red + 10 mM EDTA; 5: *Hs*prx3-6his red + 20 mM imidazole and 10 mM EDTA; 6: *Hs*prx3-6his NR; 7: *Hs*prx3-6his NR + 20mM imidazole; 8: *Hs*prx3-6his NR + 10 mM EDTA; 9: *Hs*prx3-6his NR + 20 mM imidazole and 10 mM EDTA. The reduced dimer becomes unstable with the addition of chelating agents indicating the stability of the dimer in this form is due to the presence of Ni^{2+}. The NR dimer remains stable throughout indicating the presence of an intra-dimer disulfide bond.

These results show that the protein active site is still redox sensitive. However, unlike the non-tagged construct, local unfolding of the active site that occurs upon disulfide bond formation does not lead to destabilisation at the A-type interface and the dodecamer stays intact.

4.3.3 The quaternary structure of *Hs*Prx3-WT is sensitive to the active site redox state

Within the translated protein sequence of *Hs*Prx3-6his there is a tobacco etch virus protease (TEV) recognition site. This allows for the N-terminal histidine tag and linker to be cleaved via proteolysis with rTEV protease. [29] To ensure the toroid stabilising effect was a result of the tag and not the protein handling or purification, a sample of *Hs*Prx3-6his was processed in this manner to give the WT protein + five additional amino acids:

IDPFTAPAVTQHAPYFKGTAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTE
IVAFSDKANEFHDVNCEVVAVSVDSHFSHLAWINTPRKNGGLGH*M*NIALLSDLT
KQISRDYGVLLEGSGLALRGLFIIDPNGVIKHLSVNDLPVGRSVEETLRLVKAFQY
VETHGEVCPANWTPDSPTIKPSPAASKEYFQKVNQ

The protein sequence of HsPrx3-WT post-proteolytic cleavage with rTEV. The five additional non-native N-terminal residues are in bold.

It has been shown in previous studies that the oligomeric state of *Hs*Prx3, with the tag removed, is sensitive to changes in the redox state of Cys_p with the dodecamer dissociating into dimers in non-reducing conditions. [4] Analysing the cleaved construct in this study will therefore provide a control to ensure that the stabilisation is a direct consequence of the presence of the histidine tag.

A series of different concentrations of protease to Prx were trialled at small scale to identify the optimal conditions for tag cleavage (Chapter 2 at 2.3.3). The concentration, 1:20 TEV:*Hs*Prx3, was used to cleave the tag, and the mixture was left O/N at 4^oC in the presence of a reducing agent (2 mM TCEP). At this point the cleaved and non-cleaved proteins were separated using an IMAC column that had been charged with Co^{2+} (fig. 4.6 A). The protein that was successfully digested with rTEV was not able to bind to the column. Furthermore, any uncleaved protein and the rTEV enzyme (both of which have a histidine tag) [29] would coordinate to the column and only elute upon increasing imidazole concentration. Because of this, it was the pass-through that was collected. Any additional peaks eluting upon increasing imidazole concentration were also analysed on SDS-PAGE. SDS-PAGE analysis confirmed that the flow through contained the cleaved construct (*Hs*Prx3-WT) (fig. 4.6 B).

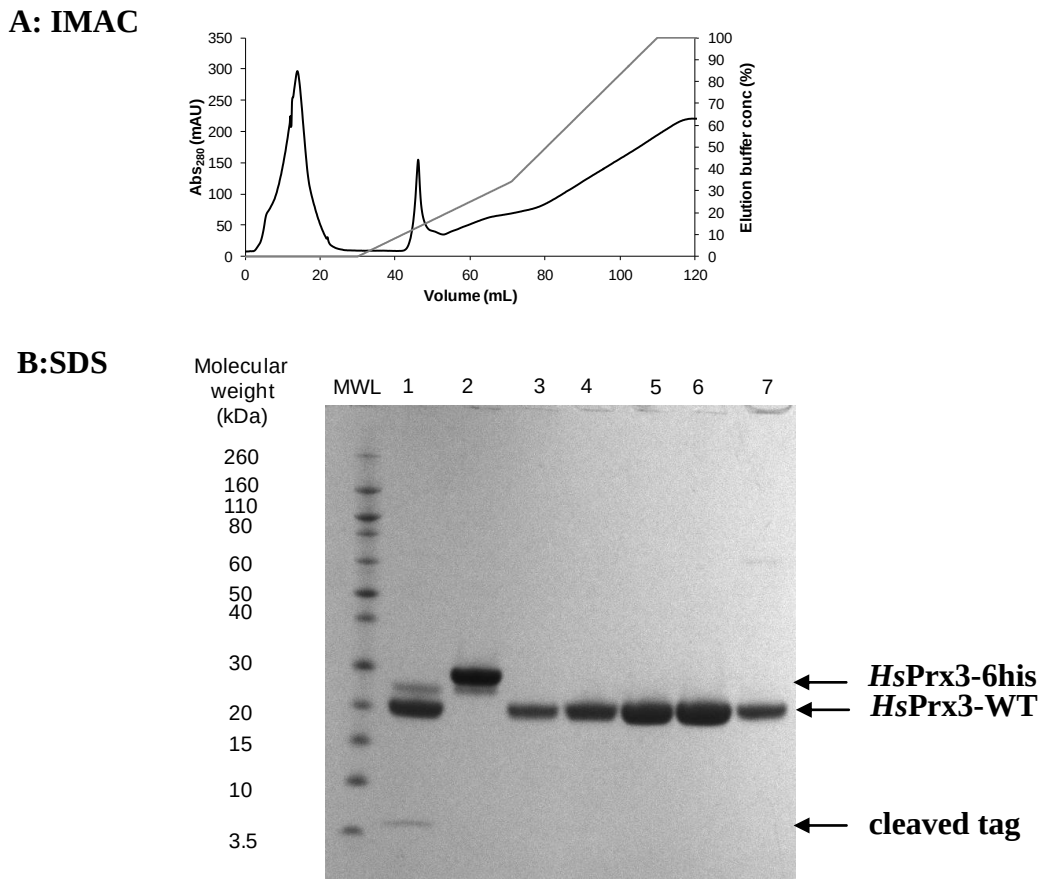


Figure 4.6: Purification of *Hs*Prx3-WT. (A) Elution of *Hs*Prx3-WT from a Co^{2+} IMAC column upon increasing imidazole concentration. Black: absorbance at 280 nm (mAU); grey: percentage concentration of 0.5 M imidazole buffer (%). (B) SDS-PAGE gel of *Hs*Prx3-WT purification. 1: *Hs*Prx3-6his + TEV protease digest crude; 2: TEV protease (~30 kDa); 4 to 6: IMAC flow through containing the isolated cleaved construct (~20 kDa); 7: weakly bound cleaved product with a small amount of uncleaved dimer present. The protein post-proteolysis has the MW of the wild type monomer.

Initial SEC analysis showed the *Hs*Prx3-WT peak eluting later than the tagged dodecamer, indicating a smaller molecular weight due to the removal of 33 residues at the N-terminus (fig. 4.7). The elution volume showed a molecular weight of 250 kDa, close to the calculated molecular weight of *Hs*Prx3-WT dodecamer species. After one hour in the absence of a reducing agent (2 mM TCEP) the dodecameric form still persisted (fig. 4.8 A). However, dodecameric rings had completely dissociated into dimers after 2 days at 4 ^{o}C (fig. 4.8 B).

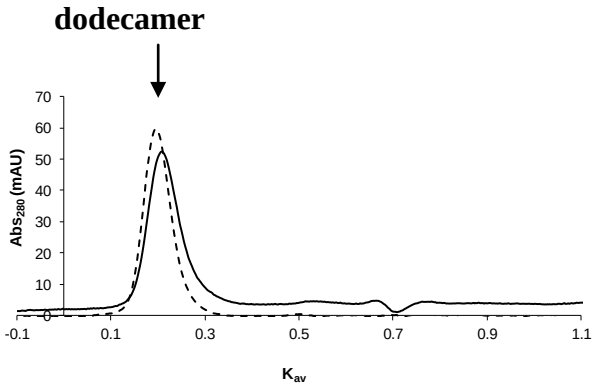


Figure 4.7: SEC of tagged and untagged constructs under reducing conditions (150 mM NaCl, 20 mM HEPES, 2 mM TCEP, pH 8.0,1 mg/mL). Black: *Hs*prx3-WT dodecamer; dotted: *Hs*Prx3-6his dodecamer. The untagged construct has a smaller molecular weight due to the loss of the tag and linker residues.

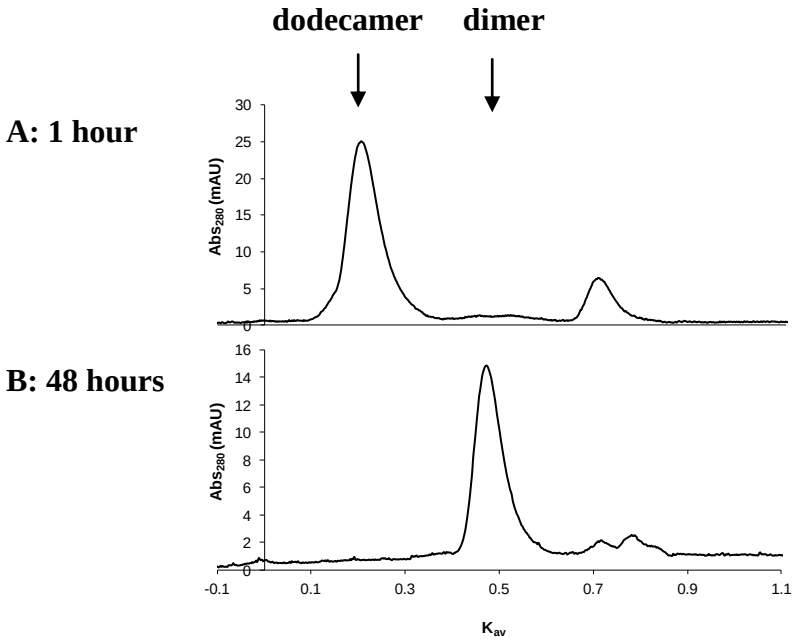


Figure 4.8: SEC of *Hs*Prx3-WT under non-reducing conditions. (A) After 1 hour and (B) after 48 hours. With time, in the absence of reducing agent, the active site disulfide bond can form leading to destabilisation at the A-type interface [2] and the complete disassembly of the dodecamer into dimers.

The method of leaving the protein over time instead of directly adding H_2O_2 was employed to avoid overoxidation. The active site sulfenic acid can undergo further oxidation by an additional H_2O_2 [1, 4, 14] to form sulfinic acid which the resolving cysteine is unable to reduce to form the intra-dimer disulfide bond. [10, 30] This not only leads to inactivation of the enzyme [10] but also restricts the local unfolding, which is believed to be needed to destabilise the A-type interface, [2, 22] and the subsequent dissociation of the toroid. [2, 31] A stable toroid in this instance would give a false positive result.

These results made it clear that the stabilisation of the toroid form of *Hs*Prx3-6his was a direct consequence of the tag, and was not due to any additional changes that may have occurred during protein expression and purification.

4.3.4 Discussion

There are a number of explanations for toroid stabilisation with the addition of a histidine tag. Histidine tags are generally accepted to have minimal effect on the structure and function of recombinant protein. [32] However, there are exceptions where the histidine tag can interact directly with the native structure, [33] reduce enzyme efficiency, [34] have a long range influence on the protein dynamics, [35] and even disrupt the protein quaternary structure. [36]

Additional amino acids on the N-terminus of other typical 2-cys Prxs have been shown to have a long range effect on the positioning of key residues at the B-type interface (Chapter 1 at 1.6.6). [19] The β-slip of 6 Å and rotation of 17.2 ^{o} around the axis parallel the β7 strand in the tagged form of *Plasmodium yoelii* Prx 1a (*Py*Prx1a^{N*}) result in changes in the ring geometry, and subsequently the tagged protein exists as an octamer rather than the native decameric form (*Py*Prx1a). [19] There is a less pronounced disparity between the dimensions of the wild type and histidine tagged toroids of *Hs*Prx3, with the tagged toroid being more compact. [4] It is possible that changes at the B-type interface are repositioning key interacting residues at the A-type interface, which in turn would stabilise the ring. A crystal structure of *Hs*Prx3-6his would be useful to confirm this hypothesis.

Prxs have a variety of different functions, one of these being their ability to behave as a putative chaperone (Chapter 1 at 1.6.7). [3, 7, 9, 14, 30, 37] The histidine tag and linker peptide used in the production of *Hs*Prx3-6his are unusually long at 33 residues; the average

length is around 14. [32] A BLAST search of protein structures that contained the same tag and linker sequence at the N-terminus was conducted and gave a number of different resolved structures. The majority of tags and linkers were disordered. It is possible that the presence of a disordered peptide in the dodecamer lumen of peroxiredoxin is initiating a chaperone response which would lead to the locked ring structure. This hypothesis will be explored in more detail in chapter 5.

Finally, histidine tags have the potential to interact with each other, with other parts of the protein, [39] and to influence the quaternary structure of recombinant proteins. [36] It is possible that the histidine tag in this construct is interacting directly with the protein, enabling the stabilisation of the toroid. However, a crystal structure would be needed to confirm this which was not within the scope of the current project.

4.4 Assembly of *Hs*Prx3-6his via divalent metal coordination

Histidine tags have been used in the field of protein science for many years as an ideal ligand to coordinate to divalent metals for the purification of recombinant proteins (Chapter 1 at 1.4.3). [40, 41] In many instances the tag is removed after purification using an enzymatic cleavage site. [32, 42] In the case of Prxs the histidine tag can not only have an effect on the structure and function of the enzyme, but can also be utilised to drive HMW oligomer formation (Chapter 1 at 1.6.8). [18, 19] There are many examples in the literature of metal coordination being utilised to drive supramolecular assembly. [43-46] Furthermore, there are numerous instances in nature where the symmetry and function of a complex system is governed by metal ion coordination. [43, 47]

The binding of metal ions is selective, with the geometry being directly affected by the pH of the solution and the selectivity of the metal. [48, 49] At pH 8, one Ni^{2+} ion binds to six imidazole side groups with octahedral symmetry. These imidazoles may come from one or a number of different histidine tags within the protein system. If imidazoles from tags on neighbouring subunits were to interact with the same metal ion this could help to stabilise and therefore drive the assembly of HMW oligomers. Due to the fact that many Prx toroids are predisposed to assemble into complex HMW oligomers, [3, 4, 6] it was predicted that the addition of metal ions would aid this process. In fact, it has been shown in the literature that chelation of Ni^{2+} to the histidine tags on *Sm*Prx1 decamers is enough to stabilise the interactions at the ring interface (R-type interface) so that the toroids assemble into tubes. [13] Furthermore, chelation to metals adds an additional route towards functionalisation of any protein nanomaterials formed from *Hs*Prx3-6his.

4.4.1 Nickel concentration influences the degree of oligomerisation

In order to investigate the optimal conditions needed for oligomerisation via metal chelation a range of metal ion concentrations were trialled, from 1:6 to 6:1 ratio of metal ion:protein. This information was used to determine the ratio of *Hs*Prx3-6his to metal ion needed to drive the formation of HMW assemblies without forming random aggregates. As an additional precaution, 20 mM imidazole was added to the buffer and removed in a step wise fashion.

This was to control the rate of metal ion to histidine coordination and therefore reduce the formation of random aggregates. A number of different methods were employed to measure the assembly of the toroids in response to metal chelation, SEC, AUC and TEM. SEC and AUC showed the increase in molecular weight and TEM was used to determine the types of structures formed.

Size Exclusion Chromatography

The protein was transferred into non reducing metal ion coordination buffer which contained 20 mM imidazole (Chapter Two at 2.7.3) and $Ni^{2+}$was spiked in at a ratio of 6:1, 3:1, 1:1, 1:3 or 1:6 metal ion to *Hs*Prx3 monomer (*Hs*Prx3-6his concentration = 1 mg/mL, ~ 40 μM). A small amount of the sample at the 1:1 ratio was analysed on SEC to ensure the protein was not aggregating in the presence of imidazole (fig. 4.9 A). The solution was left O/N at 4°C. After this time the imidazole in the buffer was removed in a stepwise fashion via diafiltration to allow for metal ion coordination.

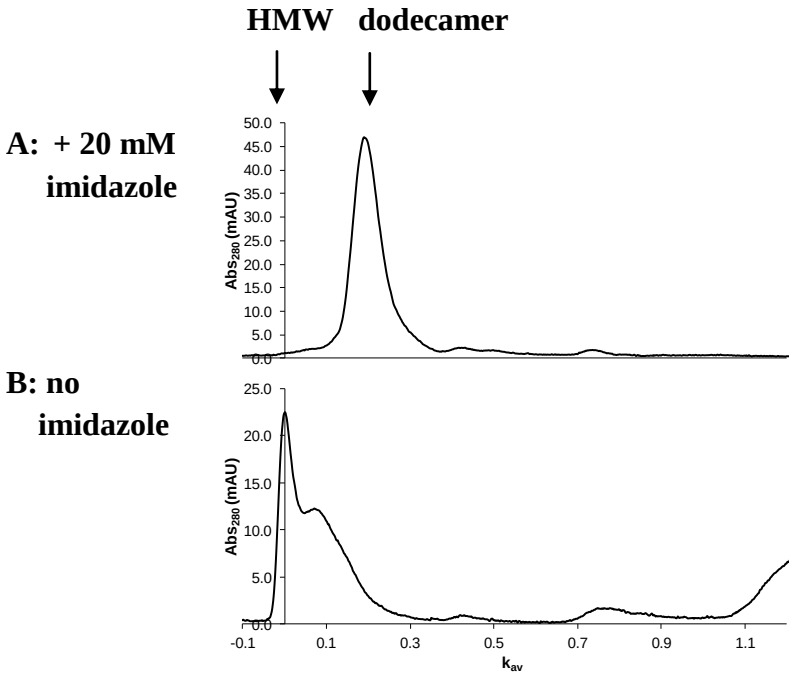


Figure 4.9: SEC trace of *Hs*Prx3-6his with 1:1Ni^{2+}:monomer (~40 μM). Both samples are at 1 mg/mL, pH 8.0 non-reduced buffer. (A) + 20 mM imidazole, (B) no imidazole. Upon removal of imidazole the protein assembles into HMW oligomers of 2 or more rings (> than the column limit of 660 kDa).

The samples containing 6:1 and 3:1 metal ion to *Hs*Prx3 precipitated out during diafiltration. It was possible to resolubilise the protein with the addition 10 mM EDTA and therefore it was believed that excess Ni^{2+} led to the disordered aggregation of the protein. Because of this, no further testing was conducted on samples containing these ratios.

There was a drop in protein concentration for the 1:1 metal ion to *Hs*Prx3-6his sample, as measured by Bradford assay (1 mg/mL – 0.6 mg/mL). [50] This was attributed to the diafiltration method used to exchange the buffer and the concentration of Ni^{2+} was adjusted accordingly to ensure the ratio remained the same. It was clear from the SEC trace (fig. 4.9 B) that the metal had a considerable effect on the oligomeric state of the protein. The peak shifted towards the column void with a large shoulder eluting at 600 kDa. This indicated the MW of stacks of two or more dodecamers.

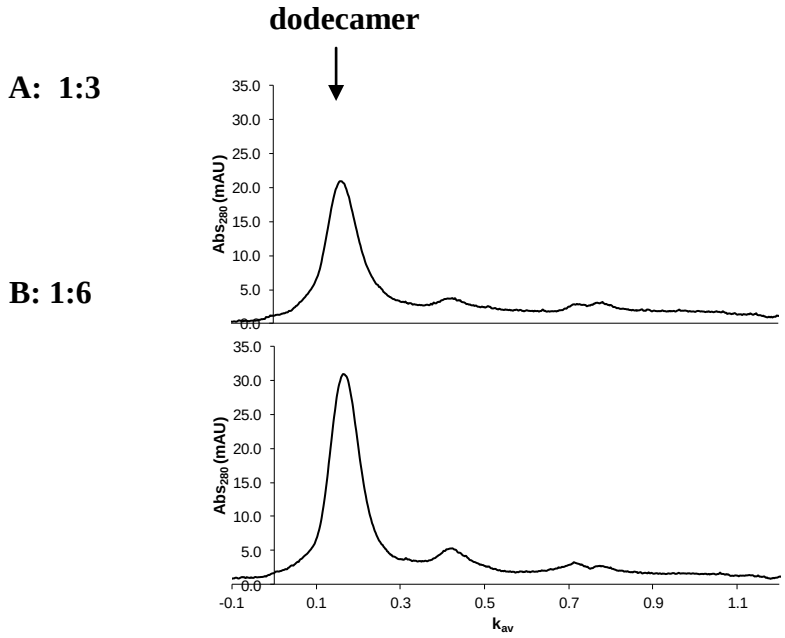


Figure 4.10: SEC trace of *Hs*Prx3-6his with (A) 1:3 and (B)1:6 Ni^{2+}:monomer ratio. In both cases there is not a sufficient concentration of Ni^{2+} to fully initiate assembly. The peak at 1:3 is slightly broader indicating a possible mix of species.

The degree of oligomerisation decreased with decreasing metal ion to *Hs*Prx3-6his ratio. The sample at 1:3 had a single peak just above the molecular weight of a single toroid. The peak was broader than was usually seen for a single dodecamer, indicating that there may be a small amount of stacking (fig. 4.10 A). At 1:6 the species eluted at the molecular weight of a

single ring (fig. 4.10 B). These results showed that the ideal concentration for the metal association was at 1:1 ratio metal ion to *Hs*prx3-6his monomer.

Analytical Ultracentrifugation and Transmission Electron Microscopy

Due to the size limits of gel filtration, the mass of the protein oligomers was measured using AUC. The protein was exchanged into a buffer containing 20 mM imidazole, and Ni^{2+} was spiked in at a 1:1 ratio with the protein monomer. Imidazole removal was carried out via diafiltration as before, and the sample was prepared for AUC analysis with a small aliquot retained for TEM.

AUC showed that the predominant species was a double stack with additional peaks at the molecular weights of stacks of three or four rings (fig. 4.11 A). Interestingly there was still a small amount of the single ring present. This peak was not seen exclusively on SEC but the broad shoulder around the double stack range did extend into elution volumes corresponding to the dodecamer molecular weight. There was also a peak relating to the dimer. This was due to the lower concentration of sample measured on AUC (0.3 mg/mL).

TEM grids were prepared following procedures as described in Chapter 2 at 2.4.4. The images corroborated with the AUC data with a mixture of single rings, double, triple and quadruple stacks (fig. 4.11 B, I - IV). It was noted that there were a number of species that resembled cages (fig. 4.11 B, V) similar to structures seen in the literature. [6] However, upon closer inspection, the parameters of these species did not match well with the diameter of a cage. It was hypothesised that the electron density visible within the lumen was in fact the tags from multiple monomers within a single toroid associating to the same Ni^{2+} ion.

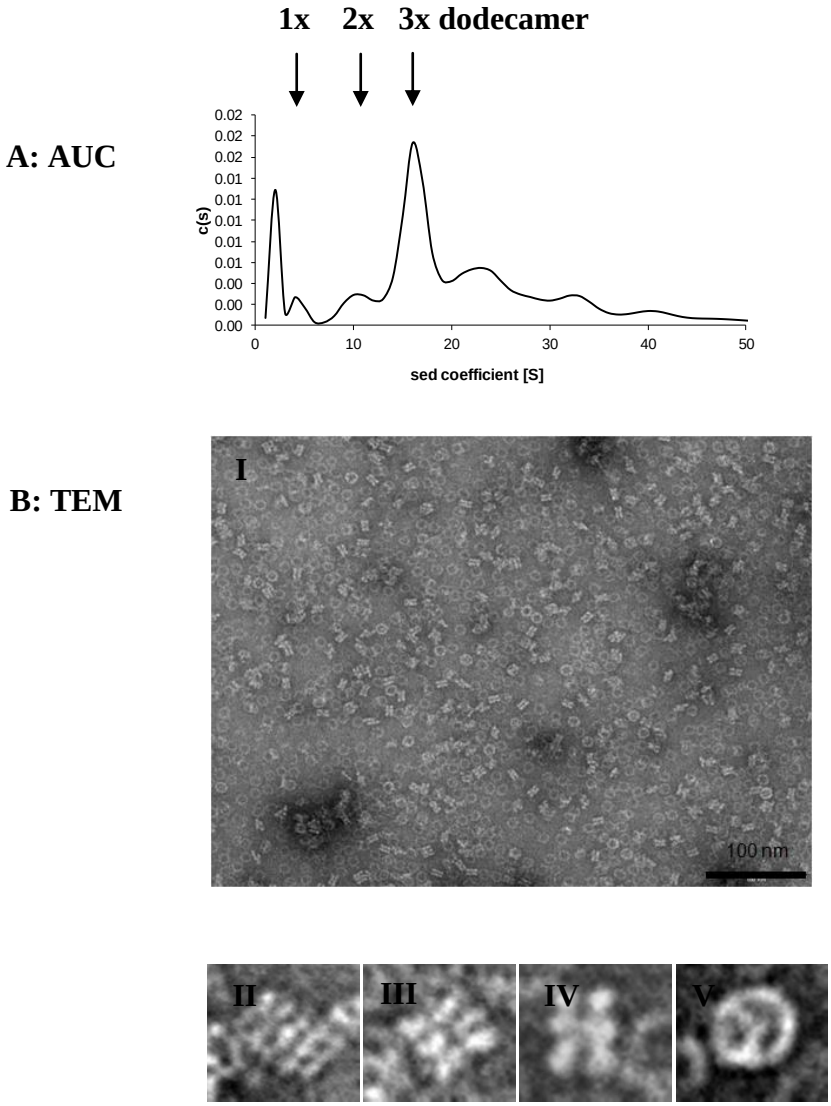


Figure 4.11: Further size and structural analysis of *Hs*Prx3-6his in the presence of equimolar amounts of Ni^{2+}. (A) AUC shows a mixture of single rings and stacks of 2 – 4 dodecamers with the most prevalent species being a double stack. (B) TEM, diluted to 50 µg/mL in dH_2O prior to grid preparation and stained with 1% uranyl acetate (140 k x magnification) I: full EM image, black bar represents 100 nm; II: stack of 5 rings; III: stack of 3 rings; IV: stack of 2 rings; V: toroid with central electron density. The addition of Ni^{2+} initiates assembly via metal coordination into HMW oligomers.

4.4.2 Disassembly with EDTA

Previous methods to remove Ni^{2+} ions that had been leached from the IMAC column showed that EDTA had sufficient chelating efficiency to disassemble the HMW species (4.2). An excess of EDTA was spiked into the sample to a final concentration of 50 mM. The solution was incubated for at least one hour at $4^{o}C$ before using dialysis to remove both the EDTA and Ni^{2+}.

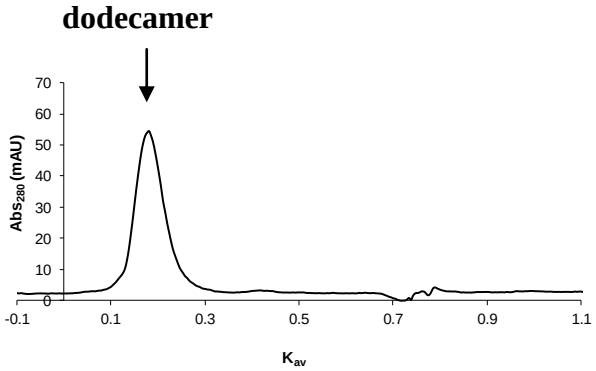


Figure 4.12: SEC trace of *Hs*Prx3-6his after the removal of metal via EDTA chelation. The chromatogram shows a single species eluting at the molecular weight of the dodecamer. Stripping the metal from the HMW species removes the means by which they can assemble. This provides a switch between the LMW and HMW oligomers.

To ensure full dissociation had occurred, the sample was analysed using SEC where a single peak eluted around the molecular weight of a dodecamer (305 kDa) (fig. 4.12). This was further proof that the formation of HMW species was a direct consequence of metal ion coordination and also provided a facile route to switch between the two forms.

4.4.3 Different divalent metals

The affinity of the imidazole side groups on the histidine tag for different divalent metals can influence the binding. [49] Recombinant proteins can be purified using metals other than Ni^{2+} on the IMAC due to over or under-protein coordination to the column. [41] This is related to the affinity and selectivity of imidazole to the divalent metal:

Cu^{2+}, Ni^{2+}, Zn^{2+}, Co^{2+}

Affinity ←

→ Selectivity

To ensure that the oligomerisation of *Hs*Prx3-6his was a direct consequence of Ni^{2+} coordination a range of different divalent metals were trialled.

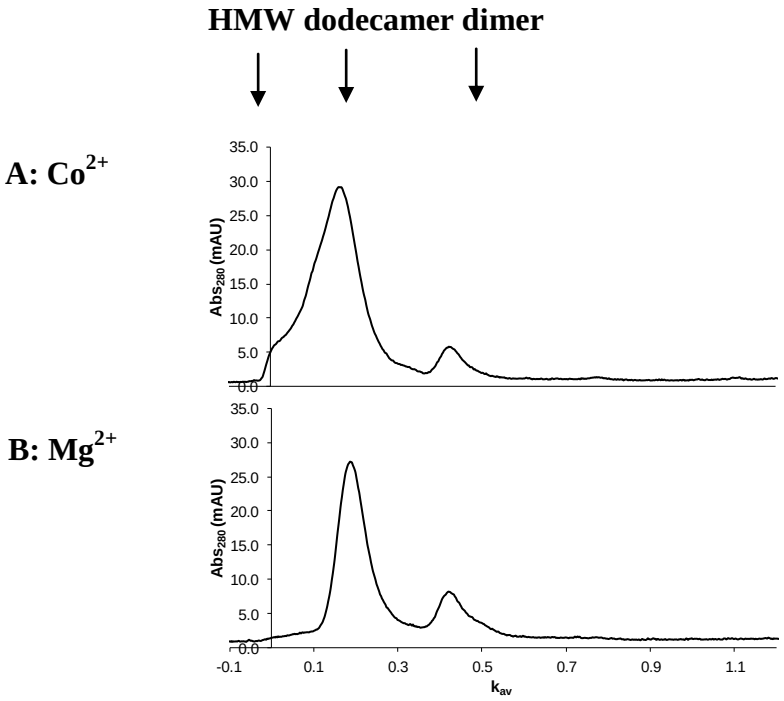


Figure 4.13: SEC of *Hs*Prx3-6his with equimolar amounts of divalent metal (~40 µM). (A) Co^{2+} and (B) Mg^{2+}. The extent of oligomerisation reduces with decreasing affinity of the metal for imidazole.

Samples were prepared as before with addition of either Mg^{2+} or Co^{2+}. Of all the divalent metals Ni^{2+} had the most notable influence, with a peak eluting at the void volume flanked by a shoulder peak (fig. 4.9 B, above). Co^{2+} instigated some stacking but a significant amount less than Ni^{2+} (fig. 4.13 A). Unsurprisingly, Mg^{2+}, which has little to no affinity with imidazole, had no effect on the oligomeric state of the protein with a single peak eluting around the molecular weight of a single dodecamer (fig. 4.13 B). This demonstrated that the oligomerisation of Prx was a direct consequence of divalent metal to histidine tag coordination as the efficiency of the system decreased with decreasing metal affinity. Furthermore, it provided a potential degree of control around the sizes of oligomers formed. This could be utilised in future work when employing this method to form protein nanomaterials for varying applications.

4.4.4 Different buffer conditions

Throughout the trials with divalent metal the degree of stacking was not to the same extent as had been seen in the literature. [13] Because of this, the conditions in which assembly occurred were varied so as to exactly match those used in the previous study. MOPS buffer was used at pH 7.6. [13] It was important to lower the pH from pH 8.0 due to the efficiency of MOPS to act as a suitable buffer above this point, and to keep the conditions equal to those reported in the literature. [13] It has been shown that histidine tags still associate with Ni^{2+} in the same geometry at this pH. [48]

Before any metal was added to the system the sample with MOPS and imidazole was analysed on AUC as a control to identify major changes to the oligomeric state that might be due to changes in buffer conditions. AUC showed that the sample was almost entirely a single ring with a small amount of HMW oligomers visible on the trace (fig. 4.14 A). This was deemed a sufficient base line to measure any additional assembly as a consequence of Ni^{2+} binding.

To induce assembly, Ni^{2+} was added in the presence of imidazole (20 mM) and the sample was left O/N at 20^{o}C. [13] After the stepwise removal of imidazole, the protein was analysed using AUC. While there was an increase in the amount of HMW oligomers, the single ring was still the most prevalent (fig. 4.14 B).

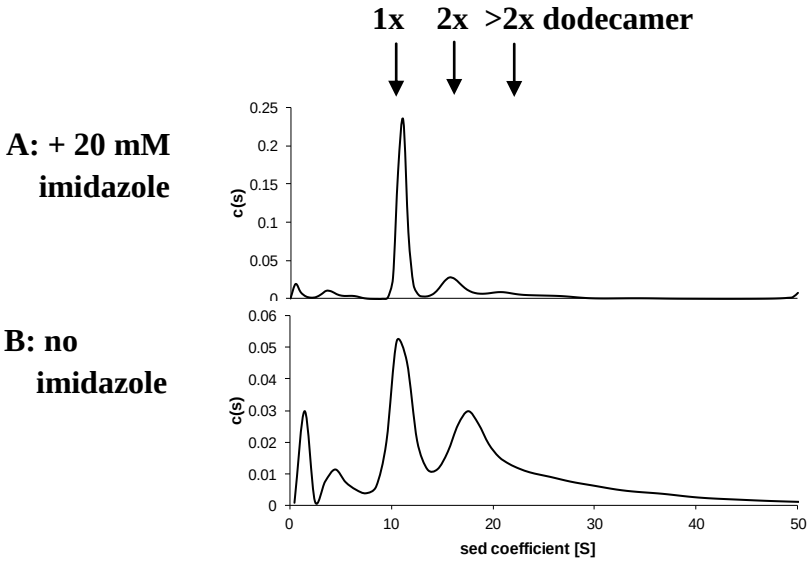


Figure 4.14: AUC of *Hs*Prx3-6his in MOPS pH 7.6 buffer with or without equimolar amounts of Ni^{2+} (40 µM). (A) Sample without Ni^{2+}, there is a small amount of oligomerisation and (B) With no imidazole the degree of assembly increases but the major oligomer is still a dodecamer.

4.4.5 Different pH conditions

In addition to the conditions described above, the influence of pH variations on the oligomerisation with metal was also investigated. The geometry of Ni^{2+} to histidine coordination can vary at pH 7.0 and below. [4, 48] This could lead to changes in the orientation of the toroids with respect to each other and add another approach towards controlling the size and shape of the protein nanomaterials.

Due to the fact that the *Hs*Prx3 can already assemble at pH conditions below 7.0, [4] pH 7.0 was chosen as the starting point. While a small amount of HMW oligomers were visible at pH 7.6 (MOPS), the assembly into larger structures increased significantly at pH 7.0 (HEPES) without the addition of Ni^{2+}. The fact that the protein oligomerised at pH 7.0 would make it difficult to discern any changes in molecular weight that were due exclusively to metal chelation. Furthermore, the degree of assembly at pH 7.0 was far greater than that achieved via Ni^{2+} binding. For this reason, no further experiments were completed on this sample and the focus was changed towards the pH sensitivity of *Hs*Prx3-6his. In future work

it may be possible to utilise the histidine tags as a form of functionalising these protein nanomaterials, formed from the pH sensitive *Hs*Prx3-6his.

4.4.6 Discussion

The use of divalent metal coordination was successful in driving the association of single *Hs*Prx3-6his toroids. This was unsurprising due to the indication of HMW structures forming as a consequence of leached Ni^{2+} from the IMAC column. However there was a significant amount of variability between the structures formed. In previous studies using a decameric typical 2-cys Prx, Ni^{2+} had been successfully utilised to aid *Sm*Prx1-6his assembly into elongated 1D nanotubes. [13] These structures were not seen in the current study despite a range of different conditions being trialled. It is important to note that there are a number of significant differences between the two Prxs which may explain why the extent of divalent metal assembly is less for *Hs*Prx3-6his:

1. **The geometry of the ring:**
   *Sm*Prx1-6his is a decamer [13] whereas *Hs*Prx3-6his is a dodecamer. This means that is it likely that there are different points of contacts at the R-type interface upon oligomerisation with a divalent metal ion.
2. **The number of tags per ring:**
   It is probable that the number of tags within the ring will have an effect on assembly. The increased number of potential Ni^{2+} binding regions may explain why the protein was more susceptible to random aggregation rather than controlled association into discrete tubes. Furthermore, having multiple 6 histidine tags within one toroid made intra-toroid rather than inter-toroid coordination more likely due the D6 symmetry with which imidazole coordinates with divalent metals.
3. **The linker between the tag and the protein:**
   While the length itself is different, which may influence the extent of inter-toroid coordination, there is also an additional region of metal binding residues as part of the linker in *Sm*PrxI-6His. This is in the form of the enteropeptidase cleavage site which contains 4 aspartic acid residues in succession. [13] Ni^{2+} can coordinate to the carboxylic acid side chain [51] which adds another mode of protein association.

Overall, divalent metal ions do influence the oligomeric state of *Hs*Prx3-6his, but more work will need to be done to optimise this manner of assembly.

4.5 *Hs*Prx3-6his is highly sensitive to variations in pH

It has been well established in previous studies that adjusting the pH can have a profound effect on the oligomeric state of Prx (Chapter 1, at 1.6.6-1.6.8). [4, 15] Wild type *Hs*Prx3 assembles into 1D tubes at pH 4.0 [52] and the toroid is stable under non-reduced conditions at pH 7.0. [4] The pH sensitivity of typical 2-cys Prxs was explored in more detail by Morais *et al*. [15] They found that in many typical 2-cys Prxs there are two conserved residues at the A-type interface, a histidine and arginine. These residues stabilise the A-type interface via electrostatic interactions when the protein is in conditions below the pK_a of the histidine side group (pH 7.37). With all this in mind, the variability of the oligomeric state of Prx due to changes in pH is a good starting point from which to predict possible assembly routes. Furthermore, as was described above, *Hs*Prx3-6his does assemble into a large oligomer at pH 7, indicating that the tag has more than just a dodecamer stabilising effect.

To be able to compare the influence of pH on *Hs*Prx3-6his to those seen for *Hs*Prx3-WT, the tagged construct was initially subjected to the same pH range, pH 9.0, 7.0, and 4.0. This was to highlight any variations that could be attributed to the presence of a histidine tag. It was also hypothesised that the reduced wild type dodecamer becomes unstable at pH 9.0 due to disruption along the A-type interface, and it was interesting to note whether the tag reduces this affect. The conditions that resulted in HMW oligomerisation were explored further to understand how the assembly occurred and therefore gain more control over oligomerisation.

4.5.1 Influence of pH on the oligomeric state of *Hs*Prx3-6his

*Hs*Prx3-6his was dialysed into three different buffers at pH 9.0, pH 7.0, or pH 4.0 (Chapter Two at 2.1.7.6) O/N at 4^{o}C. The pH was checked after this time and the sample was analysed using TEM, SEC, or both. As an additional control to ensure any large oligomers seen on TEM were formed from Prx, the cleaved construct was also dialysed into pH 4.0 buffer. It

has been well documented that the protein forms tubes under these conditions [4, 52] and this provided a baseline image that any other potential Prx tubes could be compared to.

pH 9.0

*Hs*Prx3-6his at pH 9.0 dissociated into dimers, with an additional shoulder peak eluting around the molecular weight of a tetramer (fig. 4.15). The alkaline pH, well above the pI of the protein (~6), may be enough to start denaturing the protein. If this were to occur, then the key interactions at the A-type interface would be disrupted. Additional experiments to determine the folding and stability of the protein under these conditions, such as DSF or CD, could be employed to explore the phenomenon in more detail. However, as the aim of this project was to form large structures from stable tectons, the behaviour of the protein under this particular condition was not analysed further.

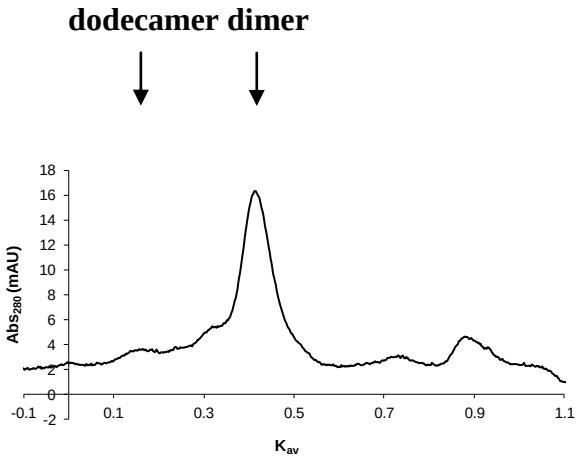


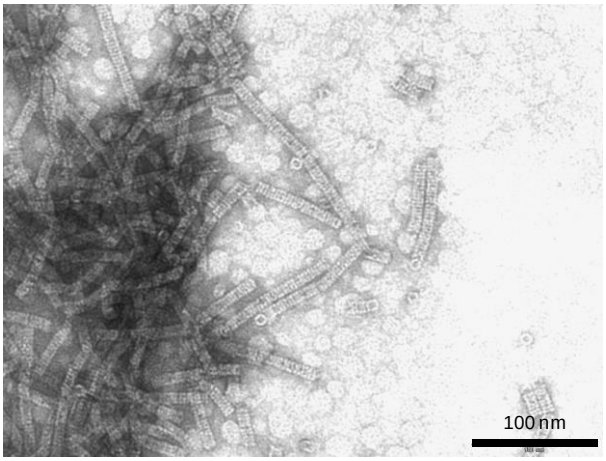
Figure 4.15: SEC of *Hs*Prx3-6his at pH 9.0. The protein has dissociated into the dimeric species with only trace amounts of the dodecamer and a shoulder peak corresponding to the tetramer.

pH 4.0

As in the cleaved construct at pH 4.0 (fig. 4.16 A) the tagged protein oligomerised into HMW assemblies. TEM analysis showed these assemblies to be 1D nanotubes (fig.4.16 B). The imidazole group on a free histidine has a pK_a of around 6. It was initially thought that the protonation of these groups at pH below 6 would create electrostatic repulsion between the

rings. However, it is clear from the TEM that this is not the case and that tube formation is not hindered by the presence of the tag. The samples at pH 4.0 were not analysed on SEC as the TEM showed tubes of multiple rings that would be larger than the size range of the column (10-600 kDa).

A: *Hs*Prx3-6his



B: *Hs*Prx3-WT

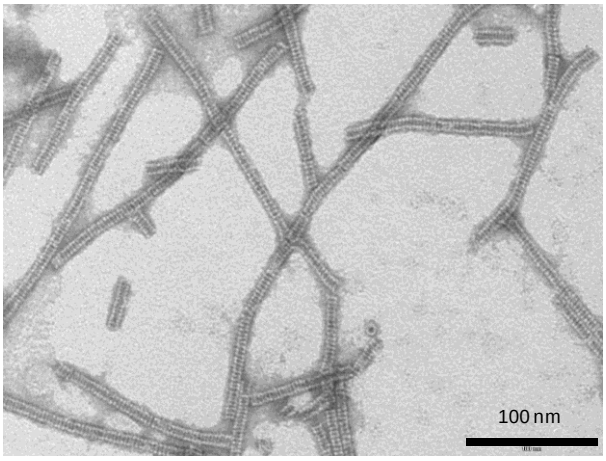


Figure 4.16: TEM of *Hs*Prx3 at pH 4.0. (A) *Hs*Prx3-6his and (B) *Hs*Prx3-WT. Both show extended tubes, with the tagged construct appearing to have slightly shorter tube lengths. The black bar represents 100 nm. The presence of the histidine residues does not hinder tube formation at pH 4.0.

pH 7.0

Perhaps the most remarkable result was the assembly of *Hs*Prx3-6his at pH 7.0. The tagged construct, when exchanged into pH 7.0 buffer, eluted entirely in the void on SEC (fig. 4.17 A). TEM analysis showed that these HMW structures were one dimensional nanotubes (fig. 4.17 B). This is the first time these particular structures have been seen for *Hs*Prx3 at physiological pH.

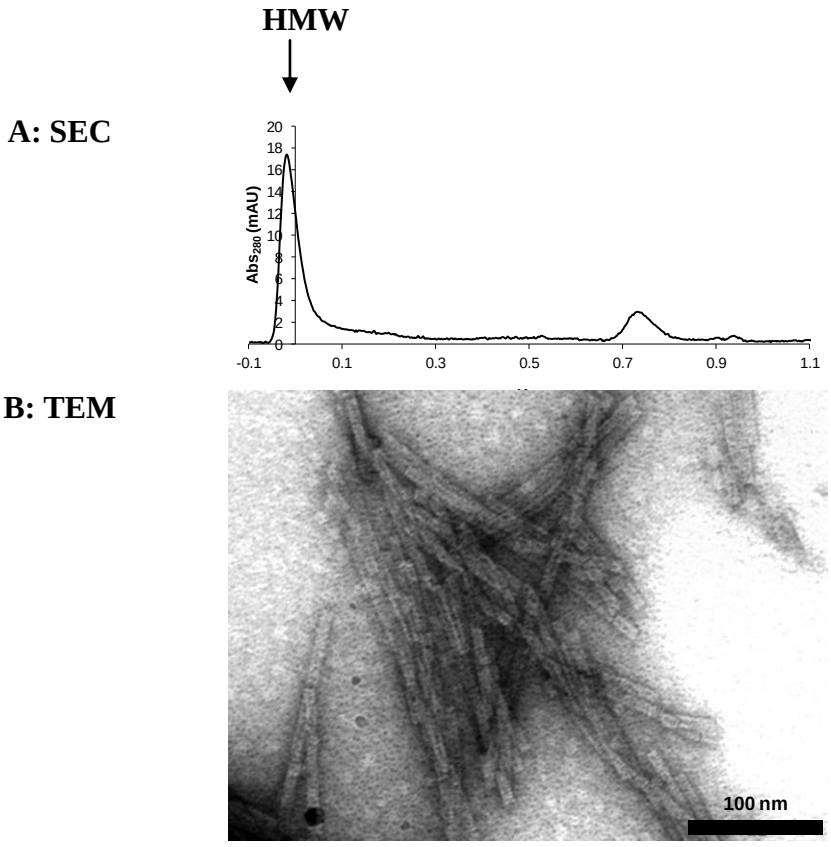


Figure 4.17: Size and structural analysis of *Hs*Prx3-6his at pH 7.0. (A) SEC trace of *Hs*Prx3-6his at pH 7.0 (1 mg/mL). All of the protein eluted at the void indicating a species greater than 660 kDa. (B) TEM of *Hs*Prx3-6his 1D tubes at pH 7.0, diluted to 50 μg/mL in dH_2O prior to grid preparation and stained with 1% uranyl acetate. The black bar represents 100 nm. Lowering the pH from 8.0 to 7.0 triggers the assembly of toroids to into long tubes.

4.5.2 pH switch

To assign a point at which the HMW oligomerisation occurred, a range of pH buffers were tested, 8.0, 7.8, 7.6, 7.4, and 7.2. (*Sample preparation, SEC and HPLC analysis were carried out by Jack Sissons, a summer student working at Callaghan Innovation between November 2014 and February 2015.)* The influence of varying pH was trialled under reducing and non-reducing conditions. This was to determine whether the structural changes that occur in the wild type enzyme during the formation of the active site disulfide bond [22] influence the formation of HMW structures.

Size Exclusion Chromatography

Non-reduced

The protein was transferred into the respective buffer by O/N dialysis. To ensure complete buffer exchange, the pH of the protein sample was measured prior to any further analysis. All SEC was carried out at 1 mg/mL to eliminate any potential differences that may be caused by fluctuations in protein concentration.

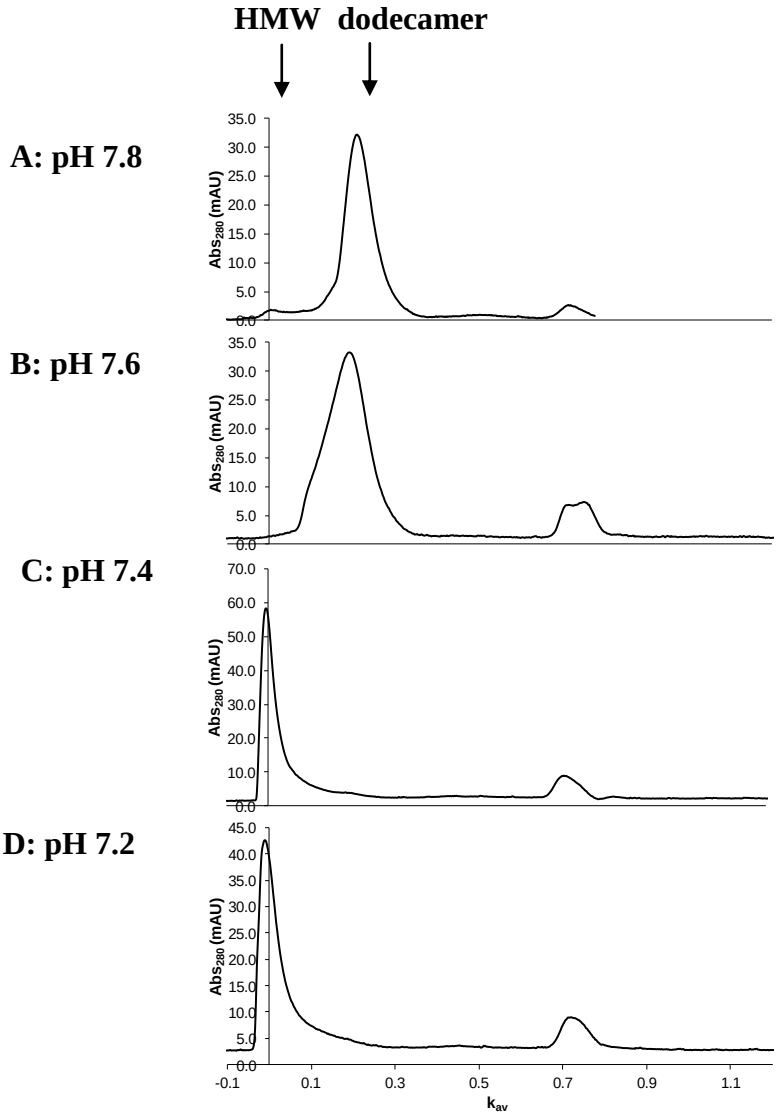


Figure 4.18: SEC chromatograms of the oligomerisation of *Hs*Prx3-6his with decreasing pH in non-reducing conditions (1mg/mL). (A) pH 7.8, (B) pH 7.6, (C) pH 7.4 and (D) pH 7.2. As the pH is lowered the protein assembles into structures with MW greater than the column limits (>660 kDa) causing the peak to shift towards the void. The switch between single rings and HMW oligomers occurs between pH 7.6 and 7.4 with a shoulder volume at HMW elution volume starting to appear at pH 7.6.

While only single rings where seen at pH 7.8 (fig. 4.18 A), the protein began to associate at pH 7.6 with a small shoulder corresponding to HMW oligomers appearing on the chromatogram (fig. 4.18 B). Between pH 7.6 and 7.4 there was a large shift towards the

column void (fig. 4.18 C) indicating structures above the molecular weight range of the column (> 660 kDa). The same profile was seen at pH 7.2 (fig. 4.18 D) which accounts for a switch between single rings and HMW structures occurring between pH 7.4 and 7.6.

Reduced

In previous studies it had been hypothesised that rearrangement of the active site that occurred upon disulfide bond formation between peroxidatic Cys_p and Cys_r, is needed to facilitate the interactions at the R-type interface that instigate oligomerisation of *Sm*Prx1 decamers into HMW assemblies. [3] If the addition of a reducing agent hindered the assembly within the current system, this would give an indication of the types of interactions at the R-type interface that were driving the assembly of the tagged construct at physiological pH.

The reduced samples were subjected to the same conditions as the non-reduced samples with the addition of 2 mM TCEP. For all of the samples there was little difference between the reduced and non-reduced forms (Appendix 4.3). The sample at pH 7.4 did display a more pronounced shoulder at the lower molecular weight range but reduction of the active site disulfide bond was not sufficient to completely disrupt the stacking (fig. 4.19).

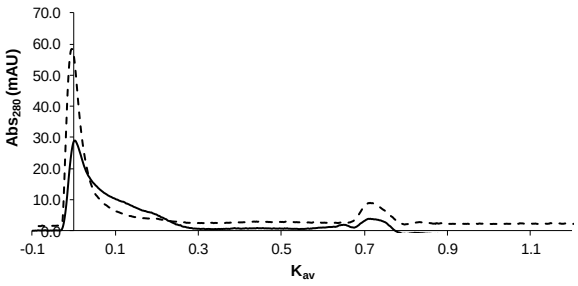


Figure 4.19: SEC trace of *Hs*prx3-6his at pH 7.4. Black: the reduced form and dotted: non-reduced. *Hs*Prx3 at pH 7.4 oligomerises to a lesser extent in the presence of reducing agent. The shoulder peak extends towards the dodecamer range when the protein is reduced.

It is important to note that there are differences in the stacking of decamers and dodecamers. In the decameric form, a rotation of the two rings with respect to each other is needed for a

cog-wheel like interaction. [3] This is seen to a lesser extent for tubes of *Hs*Prx3-WT at pH 4.0 as the geometry of the ring is different. [4, 52] As a result, while there will be similarities in the points of contact at the R-type interface between the two proteins, the interactions are not identical. This means that the local unfolding of the active site that appeared to be necessary for the formation of stacks of *Sm*Prx1 decamers [3] may not be needed for the stacking of *Hs*Prx3-6his. Furthermore, it is evident that the presence of a histidine tag is promoting assembly into tubes at a much higher pH than was seen with the cleaved construct. It is likely that there are additional points of interaction in this construct that have not been observed before.

4.5.3 Further size and structure analysis of non-reduced *Hs*Prx3-6his; pH 8.0 – 7.2

Analytical Ultracentrifugation

The samples were prepared as described previously and were run at 38,000 rpm. This velocity was chosen as it had previously given good data around the dodecamer and double stack range.

As seen on SEC, *Hs*Prx3-6his at pH 7.8 only existed as a single ring (fig. 4.20 A). The first significant change came at pH 7.6 where a small shoulder peak began to appear, reminiscent of the shoulder peak seen on SEC (fig. 4.20 B). Again, the size increased considerably between pH 7.6 and 7.4, with masses around five times the dodecamer molecular weight indicating stacks and tubes (fig. 4.20 C). Interestingly there was a further increase in size between 7.4 and 7.2 that was not visible by SEC (fig 4.20 D).

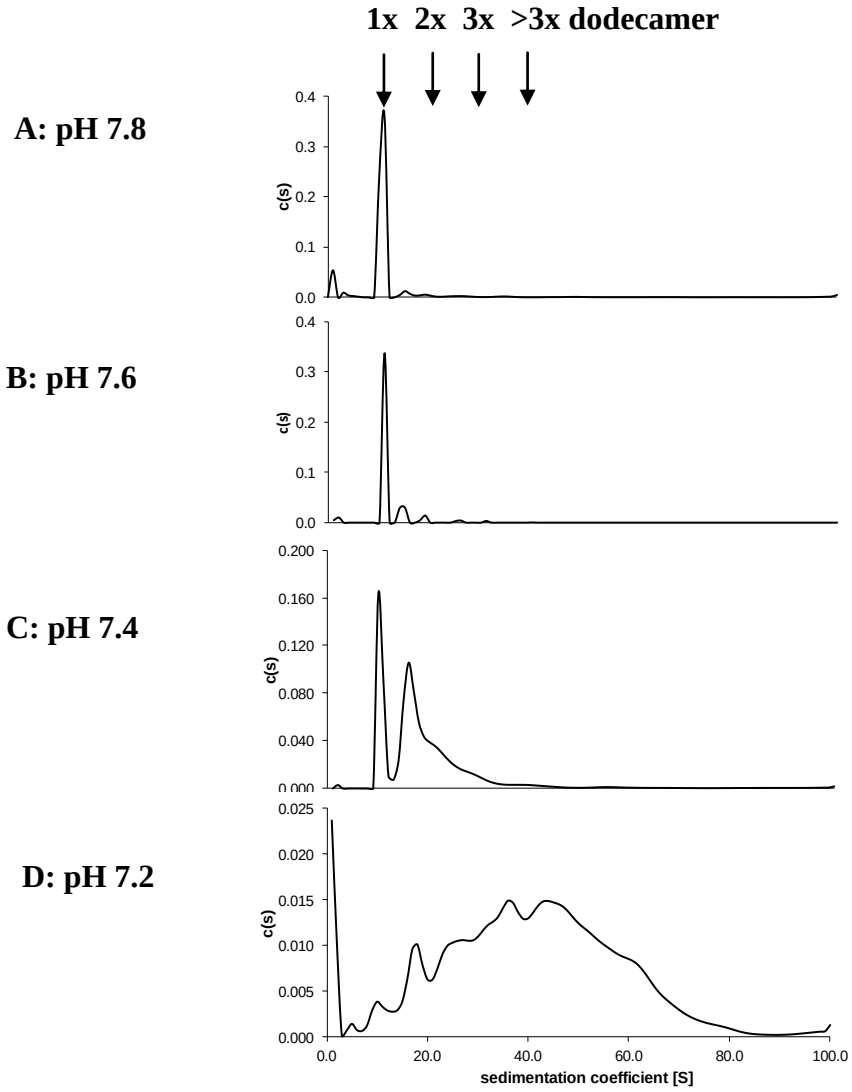


Figure 4.20: AUC traces of *Hs*Prx3-6his at varying pH. (A) pH 7.8, (B) pH 7.6, (C) pH 7.4 and (D) pH 7.2. The initial switch between LMW and HMW species occurs between pH 7.6 and 7.4. Decreasing the pH to 7.2 leads to the formation of a heterogeneous mix of large oligomers. The AUC centrifuge velocity is 38,000 rpm, which is too fast to gain enough data around the MW of the larger oligomers at pH 7.2.

Due to the large mass, the oligomers at pH 7.2 sedimented within the first few scans of the AUC run at 38,000 rpm. This meant that there was insufficient data to be able to confidently attribute masses to the peaks. Because of this, the run was repeated at a slower velocity; 20,000 rpm. At this velocity the larger oligomers sedimented over a broader range of scans which provided sufficient data around the upper end of the size ranges (fig. 4.21).

When fitting data to a heterogeneous mix of species using Sedfit, it is wise to measure the regularization by both the Maximum entropy (Equation 4.1, at 4.3.1, above) and the Tikhonov-Phillips regularization using the second derivative. [53] The latter smooths out any noise and/or artefacts from the Sedfit calculation. When this regularization was applied to the data at pH 7.2 a smooth curve was seen with a maximum around S = 55, relating to the mass of 5 x dodecamers (fig. 4.21). This shows that there was a heterogeneous mix of HMW species within the system rather than individual discrete species. However, it was still possible to discern from the data that minute changes in pH had a large effect on the oligomeric state of *Hs*Prx3-6his. This level of control has not been seen previously for Prx

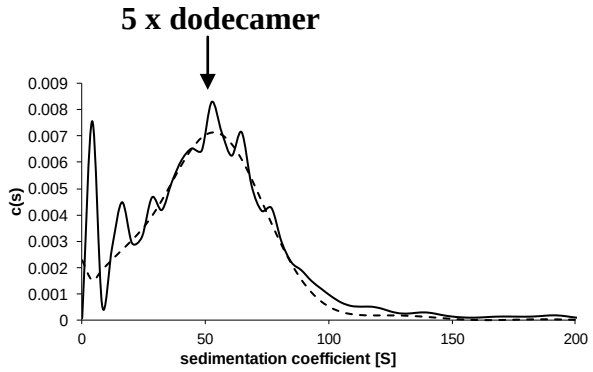


Figure 4.21: AUC trace of *Hs*Prx-6his in pH 7.2 at 20,000 rpm. When completing the regularisation using maximum entropy (black), the larger oligomers give defined peaks, consistent with discrete species, with the most abundant species being around the molecular weight of 5 x dodecamer. However, the regularisation by the 2nd derivative (dotted, with fewer artefacts) shows a heterogeneous mix of species, with the maximum at 5 x dodecamer. This is closer to a true representation of the range of molecular weights in the sample.

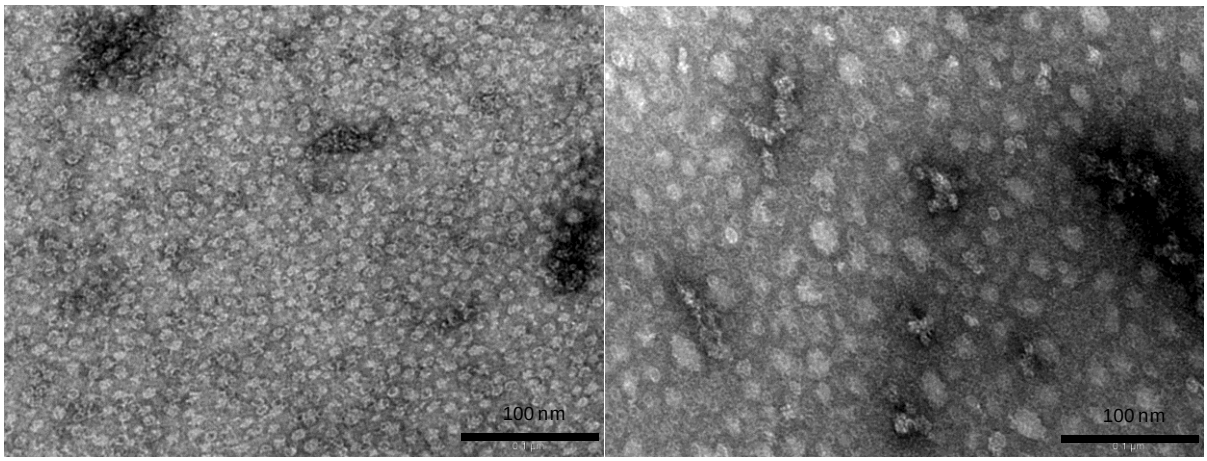
Transmission Electron Microscopy

Protein samples were transferred into the different pH buffers by O/N dialysis at 4^{o}C. The samples were diluted in dH_2O to a final concentration of 50 μg/mL immediately before TEM grid preparation. *Hs*Prx3-6his at pH 7.8 presented as single rings, similar to those seen at pH 8.0 (fig. 4.22 A). There was some evidence of the rings grouping together into clusters, although this is likely to be an artefact of the grid preparation as only a single species was

seen on AUC and SEC. At pH 7.6 stacks of two or three rings began to form (fig. 4.22 B), although there was still a significant amount of single rings present which aligns well with the profile displayed in the solution data. As expected, lowering the pH to 7.4 led to the formation of tubes ranging between 16 – 160 nm in length with the average length at 70 nm (fig. 4.22 C). Aside from a small number of outliers, the tubes were remarkably uniform in length, more so than has been seen previously. Again, this aligns with the AUC solution data. Following the trend with the solution data, there was a shift towards much longer tubes at pH 7.2 (fig. 4.22 D). The sizes ranged from 12 nm – 471 nm in length. Surprisingly, the average length was lower than that of the tubes at pH 7.4 at 56 nm. However, the standard deviation was 59 nm, meaning that the distribution of tube lengths was too great to discern an accurate average.

The TEM results agree well with the solution data with long tubes beginning to form at pH 7.4, at which point the length is particularly uniform. It is important, when designing self-assembling nanomaterials for specific function, to have control over the dimensions of the structures that are formed. At pH 7.2 the tubes extended even further, however there was less regularity between the tube sizes. It is intriguing that the pH sensitivity is over a physiological range. The same range has been noted as significant within the mitochondria,[54] the native location of *Hs*Prx3. [28] While the tagged construct is a non-native form, the interactions within the wild type sequence may be similar, with the histidine tag acting as a stabiliser to the HMW oligomers. While it was not within the scope of this project to explore, the possible physiological relevance of the pH sensitivity definitely warrants further investigation.

A: pH 7.8 **B: pH 7.6**



C: pH 7.4 **D: pH 7.2**

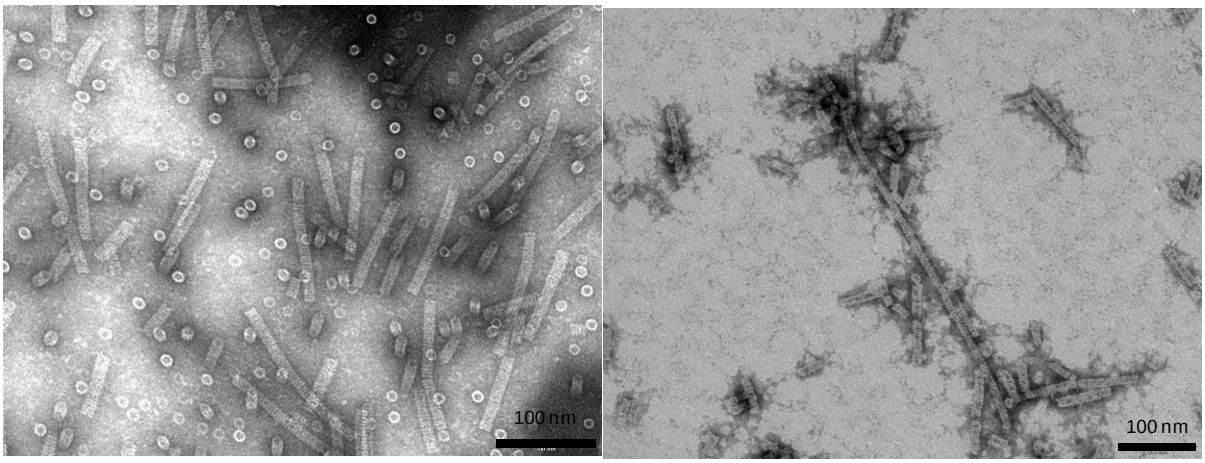


Figure 4.22: TEM of *Hs*Prx3-6his at varying pH. (A) pH 7.8, 100 μg/mL (180 k x magnification), (B) pH 7.6, 50 μg/mL (180 k x magnification), (C) pH 7.4, 50 μg/mL (140 k x magnification) and (D) pH 7.2, 50 mg/mL (110 k x magnification). The degree of oligomerisation increases with decreasing pH, with the switch been rings and stacks to tubes occurring between pH 7.6 and 7.4. This aligns with the solution data, showing that the trend towards tube formation upon lowering pH can be confirmed using TEM. The tubes at pH 7.4 are remarkably uniform, while at pH 7.2 there is a much more heterogeneous mix of lengths.

4.5.4 Impact of varying protein concentration on the degree of toroid stacking

The tecton from which the *Hs*Prx3-WT 1D nanotubes form has D6 symmetry with two orthogonal axes, a 6-fold symmetry axis perpendicular to the plane of the ring and a 2-fold symmetry axis within the plane of the ring. [52] For *Hs*Prx3-WT tubes at pH 4.0, the rings interact with a 7^{o} rotation with respect to each other which leads to the formation of structures with high order helical open symmetry. This means the rings can continue to associate with additional tectons. [55] Because of this ability to continually associate, the length of the tubes is directly influenced by the initial protein concentration. [4] Due to the similarity in the TEM images of the extended tubes of *Hs*Prx3-6his at pH 7.2 and 7.4 to those of *Hs*Prx3-WT at pH 4.0, it was hypothesised that the assembly of *Hs*Prx3-6his dodecamers would have a very similar symmetry and mode of association. If correct, this hypothesis would mean variations in the initial protein concentration of *Hs*Prx3-6his at physiological pH would lead to changes in tube length.

High pressure liquid chromatography (HPLC) was utilised to study the effect of concentration on the HMW oligomers at pH 7.8, 7.6, and 7.2 (Chapter 2 at 2.7.4). Concentrations between 0.5 and 5 mg/mL were analysed to see the influence, if any, the amount of protein within the system had on the size of the oligomers.

The sample at pH 7.8 did have a molecular weight slightly below that of a dodecamer which gradually increased as the concentration increased (Table 4.1, Appendix 4.4 A). This could be attributed to the slight concentration dependence of the toroid formation under non-reducing conditions, which results in a mixture of species within the peak. However, at no point was a resolved dimer seen. The molecular weight at pH 7.6 was higher than a single dodecamer (Table 4.1, Appendix 4.4 B). SEC, AUC, and TEM indicated a mixture of mainly single rings, and some HMW oligomers at this pH. The high molecular weight seen on HPLC could also be a mixture of species. There did not appear to be any trend between the protein concentration and the molecular weight of the peak, with the size at 0.5 mg/mL being the same as that at 5 mg/mL. However, at pH 7.2, there was a marked increase in molecular weight in response to a rise in the protein concentration (table 4.1, appendix 4.5 A).

Concentration	pH		
mg/mL	**7.8**	**7.6**	**7.2**
5	299	562	1000, 2400
3	268	581	970, 1800
1	271	571	535
0.5	269	562	380

Table 4.1: Table showing the concentration dependence of the HMW formation at varying pH. At pH 7.8, the size varies between ~ 300 and 270 kDa. It is possible that at lower concentrations, at this pH, there is an equilibrium between the dimer and dodecamer, tending towards the dodecamer, and there is a mixture of species within this peak. At pH 7.6 there appears to be a mix of HMW and dodecamer species with little difference between the concentrations. At pH 7.2 there is a clear trend towards HMW species upon increasing protein concentration, demonstrating that the tube formation is concentration sensitive.

The effect of concentration on the formation of large oligomers is intuitive. The interactions needed for the dodecamer to assemble into HMW structures are facilitated by a drop in pH. Without this pH change the interactions do not readily occur, and the higher concentration has no affect. Once the HMW oligomers can form, the concentration has a significant influence on the size of the structures. With more dodecamers within the system there is a higher probability that they will associate to form large structures.

4.5.5 R-type interface electrostatic interactions

To try and understand the mechanism of assembly and how the toroids interact, one of the low pH oligomers (pH 7.4) was subjected to a range of different salt concentrations. Along with the addition of an increased amount of sodium chloride, ammonium sulfate was also studied as it has been seen to have a detrimental effect on the formation of *Hs*Prx3-WT 1D tubes at pH 4.0. [4] To ensure that reduction in the tube length was not due to protein precipitation and therefore a reduction in concentration, controls were run on the sample at

salt concentrations 0 – 400 mM. Sodium chloride and ammonium sulfate buffers were prepared at 800 mM (50 mM HEPES, pH 8.0). This was mixed at a 1:1, 1:2 and 1:4 ratio with a sample of 2 mg/mL *Hs*Prx-6his that had been dialysed into a pH 8.0 buffer with no sodium chloride (Chapter 2 at 2.7.5.1). The final protein concentration was 1 mg/mL with salt concentrations of 0, 100, 200, or 400 mM. It was found that, while sodium chloride had little effect on the protein concentration (Appendix 4.6 A), increasing the concentration of ammonium sulfate caused the protein to precipitate over time (Appendix 4.6 B). This could be an explanation for the reduction of the tube length of *Hs*Prx3-WT at pH 4.0, and it also showed that ammonium sulfate could not be used in this experiment.

*Hs*Prx3-6his (pH 7.4, 150 mM NaCl, 1 mg/mL) was dialysed into buffer containing 300, 600, or 1200 mM NaCl. This was then analysed on SEC to measure if the salt concentration altered the size of the HMW oligomers. The protein concentration and pH of the sample were measured prior to SEC analysis to ensure that only the salt conditions had varied. For all samples, both the protein concentration and pH was stable.

The SEC trace showed that the salt began to disassemble the stacks at 300 mM sodium chloride (Table 4.2). At this concentration the peak shifted from the void towards the molecular weight of a single ring with a large shoulder. At 1200 mM NaCl, the stacks had completely disassociated into single toroids (Table 4.2, Appendix 4.7).

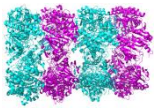
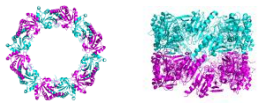
NaCl concentration (mM)	Molecular weight (kDa)	Oligomer
150	**> 600**	
300	**600** **300**	
600	**380**	
1200	**300**	

Table 4.2: Increasing the salt concentration disrupts tube formation. This suggests that assembly at the R-type interface is primarily driven by electrostatic interactions, in keeping with what has been noted in the literature.[3, 52] All samples were analysed at 1mg/mL and the concentration was checked before and after dialysis to ensure no protein had precipitated. *Images generated using Pymol with coordinates from PDB:4MH2*

Electrostatic interactions are non-covalent interactions that stabilise protein quaternary structure, and can be disrupted by a high concentration of ions within the system. [56] Sources in the literature have suggested that the association at the R-type interface is driven via electrostatic interactions. [3, 4, 52] The fact that salt does not favour the HMW oligomers suggests that, at physiological pH, the interactions that are driving the assembly of *Hs*Prx3-6his are also electrostatic.

The salt affect not only gives insight into the types of interactions that are initiating HMW assembly, but also provides a facile route towards nanostructure dissociation, thereby creating a switch between the different forms.

4.5.6 R-type interface mutations

A study conducted by Radjiana *et al.* [52] on the high resolution cryo-EM structure of *Hs*Prx3-WT tubes at pH 4.0 identified two key residues as possible points of electrostatic interaction at the R-type interface (Chapter 1 at 1.6.8). These are Lys^{23} and Asp^{24} which are located on the β-2 of the *Hs*Prx3-WT monomer (fig. 4.23). The interactions were identified by docking the crystal structure of *Sm*Prx1 chaperone species [3] into the reconstructed Cryo-EM electron density map. To explore the importance of these residues at the R-type interface, the lysine was exchanged for three different residues, K23A, K23H and K23R, listed in Table 4.3.

Amino acid substitution	Plasmid
K23	pET151 D-TOPO-hsPrx3 K23A
	pET151 D-TOPO-hsPrx3 K23H
	pET151 D-TOPO-hsPrx3 K23R

Table 4.3: List of R-type interface mutations including the plasmid into which they were cloned. The design is based on the hypothesis that lysine is a key charged residue involved in the electrostatic interactions at the R-interface.

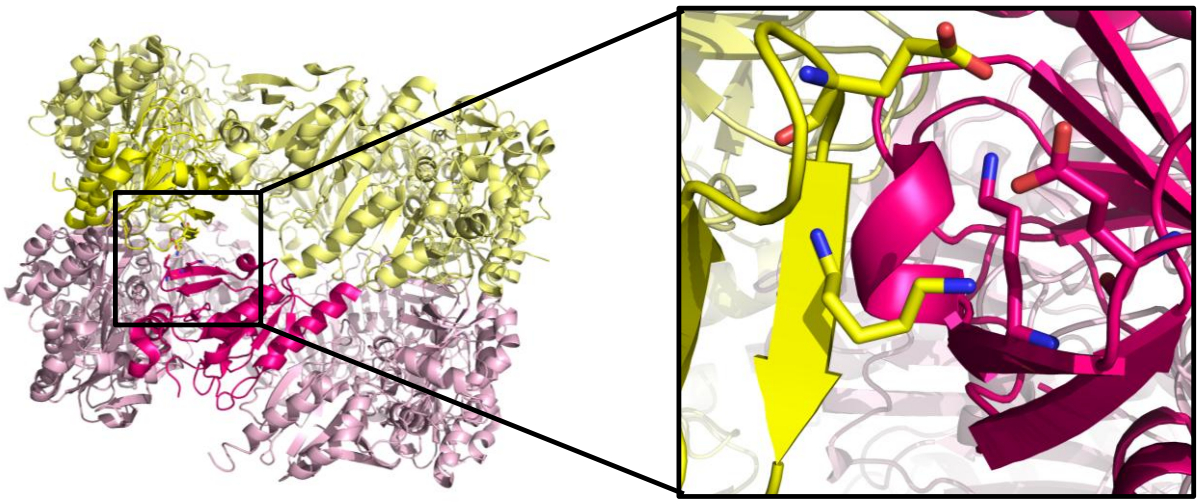


Figure 4.23: Crystal structure of *Sm*Prx1 displaying polar contacts between Asp[21] and Lys[23] at the R-type interface. [3] The image on the right shows the electrostatic interactions between Asp[21] and Lys[23] between adjacent decamers. Numbering is from *Sm*Prx1. Lys[23] on *Hs*Prx3-6his will be mutated to try and disrupt this interaction. *Images generated using Pymol with coordinates from PDB:3ZVJ.*

The plasmids where provided by Mazdak Radjiana, transformed into chemically competent BL21 Rosetta cells and expressed and purified using the same methods as *Hs*Prx3-6his. All constructs eluted as a single ring in the final SEC purification step and showed as a single band on SDS-PAGE (Appendix 4.8).

Lys[23] at the R-type interface was substituted for either histidine, arginine, or alanine. The three constructs were transferred into pH 7.2 buffer using methods described previously. SEC analysis showed that while substitution for histidine or arginine had little effect on tube assembly, substituting a charged residue for a neutral one (alanine) reduced the degree of oligomerisation (fig 4.24 A-C). However, there was still some stacking present with a broad peak tending towards the HMW range. While the results show that the substitution K23A did lead to the formation of smaller species there was still a degree of stacking. Therefore, there must be additional interactions that are influencing the HMW stabilisation.

Interestingly, it was noted that the TEM of *Hs*Prx3-K23H at pH 8.0 showed some stacks of two or three dodecamers (fig. 4.25 A), whereas only single rings were seen for the arginine substitution (fig. 4.25 B). The pK_a of a histidine side chain (6.0) is much lower than that of

arginine (12.48) or lysine (10.53). It is possible that there is a degree of charge repulsion between rings around the Lys[23] side groups at pH 8.0, and changing this for a residue with a more neutral side group is allowing stacking to occur. This was not seen in the solution studies and is likely to be attributed to the sudden increase in protein concentration upon sample drying. Nevertheless, the result shows that *Hs*Prx3-6his is predisposed to form HMW structures and that single amino acid substitutions can influence this. Furthermore, it gives more insight into the important role histidine plays in the HMW assembly.

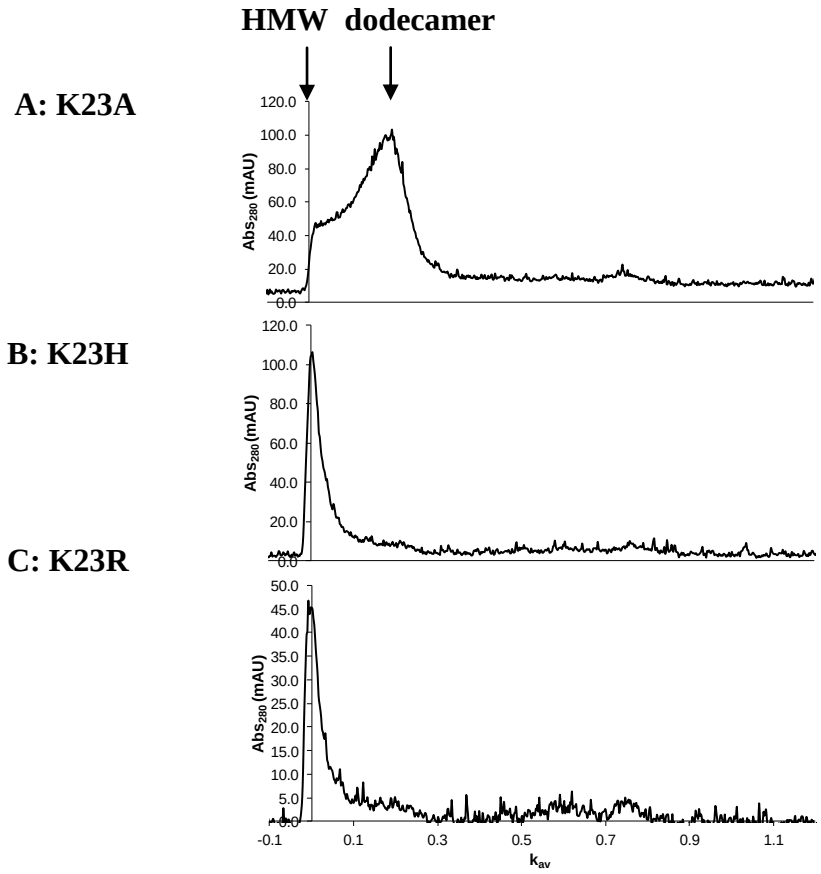


Figure 4.24: SEC of HMW oligomerisation of R-type interface K23 mutants at pH 7.2. (A) K23A, (B) K23H and (C) K23R. The oligomerisation is only partially disrupted when K23 is substituted for a neutral residue, indicating that while the lysine is involved in assembly, there are also other factors that are producing interactions at the R-type interface at pH 7.2. Substitution for a charged residue has no significant influence on the protein assembly.

A: K23H **B: K23R**

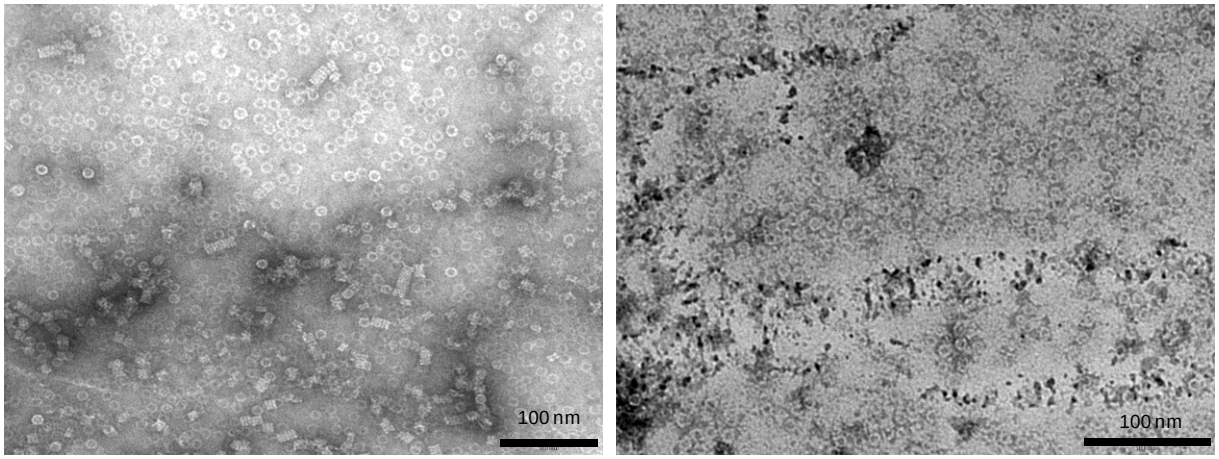


Figure 4.25: TEM of the R-interface mutantations to a charged residue (pH 8.0). (A) K23H and (B) K23R. The substitution for a histidine allows some stacking at pH 8.0.

4.5.7 Discussion

The alteration between the behaviour of *Hs*prx3-6his and *Hs*Prx3-WT has been made evident in this study. The histidine tag has an effect on the oligomeric state of the protein which, in turn, influences its response to environmental variations. The results show that the tagged protein is highly sensitive to small changes in pH, and this can be used to control the oligomerisation of the stable toroid into stacks and tubes.

There are a number of different possible explanations for the formation of tubes of *Hs*Prx3-6his at physiological pH. Firstly, histidine has been shown to be a key residue that interacts directly with an arginine at the A-type interface, which in turn stabilises the toroid under non-reducing conditions. The interaction holds the C_p loop in the locked position which stabilises the A-type interface pH values below the pK_a of this histidine (His^{113}in *Lb*Prx1m and His^{79} in *Hs*Prx3, pK_a 7.37). [15] While this particular study focused on *Lb*Prx1m, the residues involved are conserved in *Hs*Prx3. Without knowing the exact pK_a of the imidazole side chains in the histidine tag it would not be possible to say with conviction that a similar inter-toroid interaction is occurring within the current system. However, the pH switch is within the same range and definitely warrants further study. Furthermore, the sensitivity to high concentrations of salt indicated the presence of electrostatic interactions at the R-type

interface. The main additional area of charge on the protein is the histidine residues within the tag and linker sequences.

Due to the stabilisation of the A-type interface and slight variations in the geometry of the ring, [4] it is possible that there are structural changes that provide additional points of contact for the tagged construct. It has been noted that small structural changes that are not at the R-type interface can have a long range influence on the stabilisation of stacks. [57] Furthermore, additional residues on the N-terminus of Prxs have been shown to have a significant effect on the quaternary structure. [13, 17-19] Correctly placed interactions combined with the potential open helical symmetry of *Hs*Prx3-6his, make the ideal circular tecton from which stacks and tubes can form. [55] With this in mind, further structural research into the exact points of contact between the *Hs*Prx3-6his toroids would be beneficial to truly understand the mode of stacking.

Finally, it has been suggested that Prx decamers and dodecamers are predisposed to form stacks. The interactions readily form, but require changes in the structure that occur in response to external conditions in order to be stabilised. [58] If the tag and linker were to interact with the protein in a way that stabilised the stacks at pH 7.4 and below, this could allow continued assembly into long tubes. The influence of the different residues within the tag and linker sequence will be explored in more detail in Chapter 5.

Overall, the presence of the linker and tag have enabled assembly of *Hs*Prx3-6his at physiological pH. While the exact mode of oligomerisation is, as yet, unclear, the exquisite control over the assembly though small changes in pH places this protein as an ideal tecton for the formation of biological nanomaterials. With the addition of the metal binding ability of the tag, these tubes have a wealth of potential functionalities. Furthermore, with a switch that occurs within a physiological pH range, this method of assembly and disassembly could be utilised in nanomaterials for drug delivery systems.

4.6 Conclusion

Histidine tagged human Prx 3 (*Hs*Prx3-6his) proved to be an ideal candidate for the formation of switchable nanoscale biomaterials. The comprehensive knowledge afforded by the literature allowed for the prediction of model conditions needed to achieve this, and throughout this study new and tailorable methods have been developed.

The presence of the histidine tag at the N-terminus provides a simple means to stabilise the LMW dodecameric ring. This has been seen in previous research conducted on N-terminal tagged typical 2-Cys Prxs [18], but the extent to which these structural changes can be manipulated to drive assembly had not been fully explored. It is clear that while the protein differs significantly from its native form, the changes were conducive for this study.

With the addition of a histidine tag, there is the option to utilise metal coordination to form protein nanostructures. The conditions needed to bind divalent metals to histidine have been well documented since the advent of immobilised metal chromatography as a means to purify recombinant proteins. It was shown in this study that Ni^{2+} can be used to assemble Prx dodecamers into stacks. By studying different divalent metals that have a lesser affinity towards the imidazole side group, it was demonstrated that the stacking is indeed a result of metal coordination with the degree of assembly decreasing with decreasing metal affinity. Furthermore, the stacks can be dissociated with the addition of a chelating agent, providing a switch between the LMW and HMW forms.

*Hs*Prx3-WT has been shown to be sensitive to changes in pH. For this reason, a range of pH conditions were analysed to discover if this method can be used to oligomerise *Hs*Prx3-6his into tubes, and if there are any significant differences between the two homologues. Indeed, the addition of non-native residues at the N-terminus has an effect on the behaviour of *Hs*Prx3. The protein assembles into tubes at a higher pH than has been seen in the literature, with the association occurring between pH 7.6 and 7.4. In addition to this, the tubes formed at pH 7.4 are more uniform than had been seen previously. Dissociation of the tubes is achieved with the addition of relatively high concentrations of salt. This not only gives insight into the types of interactions occurring at the R-type interface that are driving assembly, but also provides a means to switch between the HMW and LMW forms. Further analysis of specific residues at the inter-dodecamer interface (that may be aiding assembly) shows that the

charged residue, lysine 23, plays a role but is not the only important residue at the R-type interface.

The capability to form switchable structures of distinct dimensions using Prx as a tecton within the physiological pH range is new. Furthermore, the degree of control in the dimensions of the materials formed is beyond what has been previously reported. It is essential, when designing nanoscale materials for a specific function, to be able to have this level of control. Therefore, the work accomplished in this study has real world significance, and has the potential to be expanded upon and utilised in the field of protein nanotechnology.

4.7 References

1. Hall A, Karplus PA, Poole LB: Typical 2-Cys peroxiredoxins - structures, mechanisms and functions. *Febs Journal* 2009, 276(9):2469-2477.

2. Wood ZA, Poole LB, Hantgan RR, Karplus PA: Dimers to doughnuts: Redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry* 2002, 41(17):5493-5504.
3. Saccoccia F, Di Micco P, Boumis G, Brunori M, Koutris I, Miele AE, Morea V, Sriratana P, Williams DL, Bellelli A *et al*: Moonlighting by different stressors: Crystal structure of the chaperone species of a 2-Cys peroxiredoxin. *Structure* 2012, 20(3):429-439.
4. Phillips AJ, Littlejohn J, Yewdall NA, Zhu T, Valery C, Pearce FG, Mitra AK, Radjainia M, Gerrard JA: Peroxiredoxin is a versatile self-assembling tecton for protein nanotechnology. *Biomacromolecules* 2014, 15(5):1871-1881.
5. Cao Z, McGow DP, Shepherd C, Lindsay JG: Improved catenated structures of bovine peroxiredoxin III F190L reveal details of ring-ring interactions and a novel conformational state. *PLOS One* 2015, 10(4).
6. Meissner U, Schroder E, Scheffler D, Martin AG, Harris JR: Formation, TEM study and 3D reconstruction of the human erythrocyte peroxiredoxin-2 dodecahedral higher-order assembly. *Micron* 2007, 38(1):29-39.
7. Karplus PA: A primer on peroxiredoxin biochemistry. *Free Radical Biology and Medicine* 2015, 80:183-190.
8. Perkins A, Nelson KJ, Parsonage D, Poole LB, Karplus PA: Peroxiredoxins: guardians against oxidative stress and modulators of peroxide signaling. *Trends in Biochemical Sciences* 2015, 40(8):435-445.
9. Banerjee M, Chakravarty D, Ballal A: Redox-dependent chaperone/peroxidase function of 2-Cys-Prx from the cyanobacterium Anabaena PCC7120: role in oxidative stress tolerance. *BMC Plant Biology* 2015, 15(60):444 - 461.
10. Cox AG, Winterbourn CC, Hampton MB: Mitochondrial peroxiredoxin involvement in antioxidant defence and redox signalling. *Biochemical Journal* 2010, 425(2):313-325.

11. Perkins A, Poole LB, Karplus PA: Tuning of peroxiredoxin catalysis for various physiological. *Biochemistry* 2014, 53(49):7693-7705.
12. Hall A, Nelson K, Poole LB, Karplus PA: Structure-based insights into the catalytic power and conformational dexterity of peroxiredoxins. *Antioxidants & Redox Signaling* 2011, 15(3):795-815.
13. Ardini M, Giansanti F, Di Leandro L, Pitari G, Cimini A, Ottaviano L, Donarelli M, Santucci S, Angelucci F, Ippoliti R: Metal-induced self-assembly of peroxiredoxin as a tool for sorting ultrasmall gold nanoparticles into one-dimensional clusters. *Nanoscale* 2014, 6(14):8052-8061.
14. Angelucci F, Saccoccia F, Ardini M, Boumis G, Brunori M, Di Leandro L, Ippolitl R, Miele AE, Natoli G, Scotti S *et al*: Switching between the alternative structures and functions of a 2-Cys peroxiredoxin, by site-directed mutagenesis. *Journal of Molecular Biology* 2013, 425(22):4556-4568.
15. Morais MAB, Giuseppe PO, Souza T, Alegria TGP, Oliveira MA, Netto LES, Murakami MT: How pH modulates the dimer-decamer interconversion of 2-Cys peroxiredoxins from the Prx1 subfamily. *Journal of Biological Chemistry* 2015, 290(13):8582-8590.
16. Nguyen JB, Pool CD, Wong CYB, Treger RS, Williams DL, Cappello M, Lea WA, Simeonov A, Vermeire JJ, Modis Y: Peroxiredoxin-1 from the human hookworm Ancylostoma ceylanicum forms a stable oxidized decamer and is covalently inhibited by conoidin A. *Chemistry & Biology* 2013, 20(8):991-1001.
17. Cao Z, Tavender TJ, Roszak AW, Cogdell RJ, Bulleid NJ: Crystals structure of reduced and of oxidized peroxiredoxin IV enzyme reveals a stable oxidized decamer and a non-disulfide-bonded intermediate in the catalytic cycle. *Journal of Biological Chemistry* 2011, 286(49):42257-42266.
18. Cao Z, Bhella D, Lindsay JG: Reconstitution of the mitochondrial PrxIII antioxidant defence pathway: General properties and factors affecting PrxII activity and oligomeric state. *Journal of Molecular Biology* 2007, 372(4):1022-1033.
19. Gretes MC, Karplus PA: Observed octameric assembly of a Plasmodium yoelii peroxiredoxin can be explained by the replacement of native "ball-and-socket" interacting residues by an affinity tag. *Protein Science* 2013, 22(10):1445-1452.
20. Cho KJ, Park Y, Khan T, Lee JH, Kim S, Seok JH, Bin Chung Y, Cho AE, Choi Y, Chang TS *et al*: Crystal structure of dimeric human peroxiredoxin-1 C83S mutant. *Bulletin of the Korean Chemical Society* 2015, 36(5):1543-1545.
21. Matsumura T, Okamoto K, Iwahara S-I, Hori H, Takahashi Y, Nishino T, Abe Y: Dimer-oligomer interconversion of wild-type and mutant rat 2-Cys peroxiredoxin. *Journal of Biological Chemistry* 2008, 283(1):284-293.
22. Perkins A, Nelson KJ, Williams JR, Parsonage D, Poole LB, Karplus PA: The sensitive balance between the fully folded and locally unfolded conformations of a model peroxiredoxin. *Biochemistry* 2013, 52(48):8708-8721.
23. Tavender TJ, Sheppard AM, Bulleid NJ: peroxiredoxin IV is an endoplasmic reticulum-localized enzyme forming oligomeric complexes in human cells. *Biochemical Journal* 2008, 411:191-199.
24. Duong-Ly KC, Gabelli SB: Gel filtration chromatography (size exclusion chromatography) of proteins. In: *Laboratory Methods in Enzymology: Protein, Pt C*. Edited by Lorsch J, vol. 541. San Diego: Elsevier Academic Press Inc; 2014: 105-114.
25. Cole JL, Lary JW, Moody T, Laue TM: Analytical Ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods in Cell Biology* 2008, 84:143-179.

26. Schuck P: Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophysical Journal* 2000, 78(3):1606-1619.
27. Cox AG, Peskin AV, Paton LN, Winterbourn CC, Hampton MB: Redox potential and Peroxide reactivity of human peroxiredoxin 3. *Biochemistry* 2009, 48(27):6495-6501.
28. Gourlay LJ, Bhella D, Kelly SM, Price NC, Lindsay JG: Structure-function analysis of recombinant substrate protein 22 kDa (SP-22) - A mitochondrial 2-Cys peroxiredoxin organized as a decameric toroid. *Journal of Biological Chemistry* 2003, 278(35):32631-32637.
29. Blommel PG, Fox BG: A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expression and Purification* 2007, 55(1):53-68.
30. Konig J, Galliardt H, Jutte P, Schaper S, Dittmann L, Dietz KJ: The conformational bases for the two functionalities of 2-cysteine peroxiredoxins as peroxidase and chaperone. *Journal of Experimental Botany* 2013, 64(11):3483-3497.
31. Wood ZA, Poole LB, Karplus PA: Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science* 2003, 300(5619):650-653.
32. Carson M, Johnson DH, McDonald H, Brouillette C, DeLucas LJ: His-tag impact on structure. *Acta Crystallographica Section D-Biological Crystallography* 2007, 63:295-301.
33. Majorek KA, Kuhn ML, Chruszcz M, Anderson WF, Minor W: Double trouble-Buffer selection and His-tag presence may be responsible for nonreproducibility of biomedical experiments. *Protein Science* 2014, 23(10):1359-1368.
34. Chen Z, Li Y, Yuan Q: Study the effect of His-tag on chondroitinase ABC I based on characterization of enzyme. *International Journal of Biological Macromolecules* 2015, 78:96-101.
35. Thielges MC, Chung JK, Axup JY, Fayer MD: Influence of histidine tag attachment on picosecond protein dynamics. *Biochemistry* 2011, 50(25):5799-5805.
36. Ben Hlima H, Ayadi D, Aghajari N, Bejar S: Differential properties of native and tagged or untagged recombinant glucose isomerases of Streptomyces sp SK and possible implication of the glycosylation. *Journal of Molecular Catalysis B: Enzymatic* 2013, 94:82-87.
37. Sharapov MG, Ravin VK, Novoselov VI: Peroxiredoxins as multifunctional enzymes. *Molecular Biology* 2014, 48(4):520-545.
38. Wang Y, Boudreaux DM, Estrada DF, Egan CW, St Jeor SC, De Guzman RN: NMR structure of the N-terminal coiled coil domain of the Andes hantavirus nucleocapsid protein. *Journal of Biological Chemistry* 2008, 283(42):28297-28304.
39. Chant A, Kraemer-Pecore CM, Watkin R, Kneale GG: Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA causes a conformational change at the DNA-binding site. *Protein Expression and Purification* 2005, 39(2):152-159.
40. Porath J, Olin B: Immobilized metal affinity adsorption and immobilized metal affinity chromatography of biomaterials. Serum protein affinities for gel-immobilized iron and nickel ions. *Biochemistry* 1983, 22(7):1621-1630.
41. Bornhorst JA, Falke JJ: Purification of proteins using polyhistidine affinity tags. *Methods in Enzymology* 2000, 326:245-254.
42. Arnau J, Lauritzen C, Petersen GE, Pedersen J: Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification* 2006, 48(1):1-13.

43. Brodin JD, Ambroggio XI, Tang C, Parent KN, Baker TS, Tezcan FA: Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* 2012, 4(5):375-382.
44. Brodin JD, Carr JR, Sontz PA, Tezcan FA: Exceptionally stable, redox-active supramolecular protein assemblies with emergent properties. *Proceedings of the National Academy of Sciences of the United States of America* 2014, 111(8):2897-2902.
45. Sapsford KE, Pons T, Medintz IL, Higashiya S, Brunel FM, Dawson PE, Mattoussi H: Kinetics of metal-affinity driven self-assembly between proteins or peptides and CdSe-ZnS quantum dots. *Journal of Physical Chemistry C* 2007, 111(31):11528-11538.
46. Zhang W, Luo Q, Miao L, Hou C, Bai Y, Dong Z, Xu J, Liu J: Self-assembly of glutathione S-transferase into nanowires. *Nanoscale* 2012, 4(19):5847-5851.
47. Bai YS, Luo Q, Zhang W, Miao L, Xu JY, Li HB, Liu JQ: Highly ordered protein nanorings designed by accurate control of Glutathione S-Transferase self-assembly. *Journal of the American Chemical Society* 2013, 135(30):10966-10969.
48. Valenti LE, De Pauli CP, Giacomelli CE: The binding of Ni(II) ions to hexahistidine as a model system of the interaction between nickel and His-tagged proteins. *Journal of Inorganic Biochemistry* 2006, 100(2):192-200.
49. Svanedal I, Boija S, Almesaker A, Persson G, Andersson F, Hedenstrom E, Bylund D, Norgren M, Edlund H: Metal ion coordination, conditional stability constants, and solution behavior of chelating surfactant metal complexes. *Langmuir* 2014, 30(16):4605-4612.
50. Bradford MM: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 1976, 72(1-2):248-254.
51. Kallay C, Sovago I, Varnagy K: Nickel(II) complexes of oligopeptides containing aspartyl and glutamyl residues. Potentiometric and spectroscopic studies. *Polyhedron* 2007, 26(4):811-817.
52. Radjainia M, Venugopal H, Desfosses A, Phillips AJ, Yewdall NA, Hampton MB, Gerrard JA, Mitra AK: Cryo-electron microscopy structure of human peroxiredoxin-3 filament reveals the assembly of a putative chaperone. *Structure* 2015, 23(5):912-920.
53. Schuck P: On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Analytical Biochemistry* 2003, 320(1):104-124.
54. Santo-Domingo J, Demaurex N: The renaissance of mitochondrial pH. *Journal of General Physiology* 2012, 139(6):415-423.
55. Angelucci F, Bellelli A, Ardini M, Ippoliti R, Saccoccia F, Morea V: One ring (or two) to hold them all - on the structure and function of protein nanotubes. *The Federation of European Biochemical Societies Journal* 2015, 282(15):2827-2845.
56. Watanabe H, Matsumaru H, Ooishi A, Feng YW, Odahara T, Suto K, Honda S: Optimizing pH Response of Affinity between Protein G and IgG Fc. How electrostatic moddulation affect protein-protein interactions. *Journal of Biological Chemistry* 2009, 284(18):12373-12383.
57. Saccoccia F, Angelucci F, Boumis G, Desiato G, Miele AE, Bellelli A: Selenocysteine robustness versus cysteine versatility: a hypothesis on the evolution of the moonlighting behaviour of peroxiredoxins. *Biochemical Society Transactions* 2014, 42(6):1768-1772.
58. Noichri Y, Palais G, Ruby V, D'Autreaux B, Delaunay-Moisan A, Nystrom T, Molin M, Toledano MB: In vivo parameters influencing 2-Cys Prx oligomerization: The role of enzyme sulfinylation. *Redox Biology* 2015, 6:326-333.

Chapter Five: Tag and linker peptides influence the pH sensitivity and oligomerisation of *Hs*Prx3

5.1 Introduction

Proteins have evolved to organise themselves into a staggering array of intricate assemblies that are directly related to their function and behaviour *in vivo*. [1, 2] The route towards this structure and function relationship has been well studied, with the oligomerisation being controlled by both the amino acid content of the protein and the environment in which it exists. [1, 3, 4] For this reason, proteins are an attractive target in the field of materials science. [5, 6] By understanding the influence of the environmental factors on a native protein, and using this to design specific differences in the amino acid content, it is possible to manipulate this innate self-assembling ability to drive the bottom-up formation of nanoscale materials with a wealth of potential functionalisation opportunities. [5, 7]

In the previous chapter, the influence of the recombinant histidine tag and linker peptide on human peroxiredoxin 3 (*Hs*Prx3) was explored. It was found that the presence of these additional amino acids on the N-terminus of the protein not only helped to stabilise an ideal tecton, dodecamer toroid, but also had a marked effect on the behaviour of the protein in response to external factors that differed from the native form. It was hypothesised that changes to the quaternary structure were related to the stabilisation of the A-type interface as this phenomenon had been noted in previous studies. [8, 9] The additional pH influence, a new route towards assembly, was attributed to electrostatic interactions at the R-type interface that helped to drive assembly into one dimensional (1D) tubes.

Histidine tags have been shown to have an influence on the quaternary structure and function of recombinant proteins, whether this be due to direct contact with the protein at key subunit interfaces, [10] interactions with the substrate binding site, [11] or changing the protein dynamics. [12] Furthermore, the presence of the charged side group on histidine has been utilised to develop a non-native pH response in the assembly of proteins. [13] It has also been identified as a residue that is involved in controlling the pH sensitivity of peroxiredoxin oligomeric state. [14] The histidine residues in the tag and linker peptide are the prime

candidates for the origin of the electrostatic interactions. With the assembly being driven by electrostatic interactions, having a large area of surface charge is likely to have an effect. Substitutions for residues that were implied in the stacking of the wild types construct at pH 4.0, did not completely disrupt the interactions at the R-type interface, further suggesting that additional residues were involved.

*Hs*Prx3 has been reported to be a putative chaperone. [15] This type of chaperone function, holdase, prevents the aggregation or misfolding of proteins but it is not involved in the ATP-dependent refolding step. [16] Peroxiredoxins can adopt this function under stress conditions such as heat or high H_2O_2 concentration. [17, 18] The toroid (single or stacked) holds unfolded or partially unfolded protein within the lumen via series of hydrophobic interactions. [19] With the unusually long linker sequence of disordered peptide at the N-terminus of *Hs*Prx3-6his, it is possible that the R-type interface stabilisation is, in part, related to a chaperone response, increasing the surface area of the toroid lumen (and thereby increasing the points of contact for hydrophobic interactions). A series of constructs with varying linker length and amino acid content were developed to probe this hypothesis.

Exploring the origins of the pH sensitivity of *Hs*Prx3-6his and its propensity to oligomerise into large structures was also important for the design of protein nanomaterials that self-assemble from peroxiredoxin in a controllable manner. In addition to the variations in the tag and linker amino acid content, the kinetics of assembly will also be probed. Producing protein nanomaterials with in-built function is also an attractive prospect in the field of bionanotechnology. Therefore the primary function of the enzyme, its peroxidatic activity, will also be examined. With the understanding that this research affords, there will be systems put in place for the further development of *Hs*Prx3 as a protein tecton.

5.1 Design of tag and linker variations

There are distinct differences between histidine tagged and cleaved *Hs*Prx3-WT, the most notable being the stability of the toroid under non-reducing conditions at pH 8. This has also been seen for bovine peroxiredoxin 3, [9] which shares a 93% sequence homology to *Hs*Prx3. *Hs*Prx3-6his is also responsive to small variations in pH, with a conversion from single rings to stacks and then to tubes at pH 8.0, 7.6, and 7.4 respectively. Unlike the dodecamer stability, this was not seen in the bovine construct. [9] The main point of difference between

the two proteins is the linker between the tag and the protein and therefore a series of constructs were designed whereby the amino acid content in the tag and linker sequence was varied to try and allocate where the pH effect was originating. The influence of varying the number of metal binding residues was also explored.

It was established in the previous chapter that the interactions at the R-type interface were electrostatic in nature, with the tubes dissociating at high salt concentration. A key residue (Lys^{23}), was identified at this interface, as it was hypothesised that it was capable of forming a salt-bridge with an aspartic acid (Asp^{24}) on the neighbouring toroid (Chapter 1 at 1.6.6). [15] While mutating this residue to alanine, and thereby removing the salt-bridge, led to a small variation in the size of the tubes and stacks (and partially disrupting the interface) there were alternative contacts involved. The poly-6-his tag represents an area of potentially charged residues within the pH range. It is possible that theses residues form electrostatic interactions with the protein/with each other to stabilise the R-type interface and facilitate the assembly into long tubes. It was identified as the first point for sequence variations, the hypothesis being that reducing the amount of charge-charge interactions from this point would reduce the pH sensitivity of the oligomerisation. The pK_a value of the histidine side chain can vary along the tag. The charge can migrate from one chain to another, ranging from acidic to basic from the N-terminal to the C-terminal end respectively. [20] Therefore changing the length of the tag also provides the possibility of varying the pK_a values of the histidine side chains within the tag.

With the differing number of histidines within the tag there is also the potential for variations in metal ion coordination. In the previous chapter it was noted that having multiples of six histidine residues within the D6 symmetry of the ring may have been detrimental to inter-toroid coordination. Changing the tag length may help to drive metal ion mediated inter-toroid stabilisation without the formation of random aggregate. The use of metal ions to assemble proteins not only overcomes the problems associated with the large surface area needed for non-covalent interactions between protein units, [21] but can also be used to direct the shape of the oligomers via the geometry of the metal to ligand coordination.[21-23]

The length of the linker is a major point of difference between *Bt*Prx3 and *Hs*Prx-6his. Non-native amino acids have been implemented in structural changes of other peroxiredoxins, [8] so it is possible this explains the difference in behaviour between the two constructs. Furthermore, many peroxiredoxins have the capacity to function as a chaperone in response

to cell stress conditions such as heat and high H_2O_2 concentration that can cause proteins to unfold. [17] With the unusually long length of the tag and linker sequence used in this study (33 aa) it is possible that *Hs*Prx3 is identifying the sequence as an unfolded protein which, in turn, may instigate a chaperone response. With the long flexible linker there is also the potential for the tag to interact directly with the protein and therefore changing the length may trigger variations in the mechanism of assembly. The linker length was therefore varied systematically by removing different parts of the sequence that were related to its function, for example the V5 epitope tag or the TEV recognition site. One construct had all of the linker residues removed, leaving the histidine tag directly at the N-terminus of the native protein sequence. A further construct was designed whereby an enteropeptidase recognition site, as described in chapter 3 (at 3.4) was placed in between the tag and the protein. It was predicted that this would have a dual function of investigating the impact of adding sequential acidic residues within the tag (DDDDKD), and also provide the means to cleave the tag completely by enteropeptidase proteolysis (leaving the "true" wild type sequence).

The sequences for all the variations are listed below with their predicted properties, calculated from the amino acid sequence (ProtParam) listed in Table 5.1.

Name	pI	MW (Da)	Aliphatic index	Hydropathicity	Instability index
HsPrx3NoTEV	6.2	23897	89.3	-0.133	24.9
HsPrx3NoV5	6.0	23919	86.5	-0.144	26.1
HsPrx3NoLink	6.2	22494	87.8	-0.132	26.5
HsPrx3Ent	5.9	23489	84.4	-0.254	25.5
HsPrx3His2	5.7	24774	89.6	-0.090	25.7
HsPrx3His4	5.8	25048	88.8	-0.177	25.5

Table 5.1: Protein parameters for the original histidine tag and linker variation constructs. The proteins were designed to probe the origin of both the dodecamer stability and pH sensitivity of *Hs*Prx3-6his. All constructs were deemed to be stable under base line conditions, pH 8.0 aqueous buffer. The main differences between the proteins are the molecular weight (MW), isoelectric point (pI) and hydropathicity.

> *Hs*Prx3NoTEV
MHHHHHHGKPIPNPLLGLDSTAPAVTQHAPYFKGTAVVNGEFKDLSLDDFKG
KYLVLFFYPLDFTFVCPTEIVAFSDKANEFHDVNCEVVAVSVDSHFSHLAWINTP
RKNGGLGHMNIALLSDLTKQISRDYGVLLEGSGLALRGLFIIDPNGVIKHLSVNDL
PVGRSVEETLRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYFQKVNQ

> *Hs*Prx3NoV5
MHHHHHHENLYFQGIDPFTAPAVTQHAPYFKGTAVVNGEFKDLSLDDFKGKY
LVLFFYPLDFTFVCPTEIVAFSDKANEFHDVNCEVVAVSVDSHFSHLAWINTPRK
NGGLGHMNIALLSDLTKQISRDYGVLLEGSGLALRGLFIIDPNGVIKHLSVNDLPV
GRSVEETLRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYFQKVNQ

> *Hs*Prx3NoLink
MHHHHHHAPAVTQHAPYFKGTAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFV
CPTEIVAFSDKANEFHDVNCEVVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLS
DLTKQISRDYGVLLEGSGLALRGLFIIDPNGVIKHLSVNDLPVGRSVEETLRLVKA
FQYVETHGEVCPANWTPDSPTIKPSPAASKEYFQKVNQ

> *Hs*Prx3Ent
MHHHHHHDYKDDDDKAPAVTQHAPYFKGTAVVNGEFKDLSLDDFKGKYLVL
FFYPLDFTFVCPTEIVAFSDKANEFHDVNCEVVAVSVDSHFSHLAWINTPRKNGG
LGHMNIALLSDLTKQISRDYGVLLEGSGLALRGLFIIDPNGVIKHLSVNDLPVGRS
VEETLRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYFQKVNQ

> *Hs*Prx3-2his
MHHGKPIPNPLLGLDSTENLYFQGIDPFTAPAVTQHAPYFKGTAVVNGEFKDL
SLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKANEFHDVNCEVVAVSVDSHFSH
LAWINTPRKNGGLGHMNIALLSDLTKQISRDYGVLLEGSGLALRGLFIIDPNGVIK
HLSVNDLPVGRSVEETLRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYF
QKVNQ

>*Hs*Prx3-4his
MHHHHGKPIPNPLLGLDSTENLYFQGIDPFTAPAVTQHAPYFKGTAVVNGEFK
DLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKANEFHDVNCEVVAVSVDSHF
SHLAWINTPRKNGGLGHMNIALLSDLTKQISRDYGVLLEGSGLALRGLFIIDPNGV
IKHLSVNDLPVGRSVEETLRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKE
YFQKVNQ

New construct sequences. The tag and linker residues are in bold.

During the write up of this thesis it became apparent that the plasmids that were provided by Epoch (Epoch Life Science, Inc, US) did not actual contain the requested sequence. The plasmid map that was provided by the company showed the protein sequence, but no DNA sequencing data was provided. When this was sent upon request, it showed that there was an additional 6-histidine tag and thrombin cleavage site, translated at the N-terminus of the protein due to the restriction sites (between the NdeI and XhoI) into which the open reading

frame had been cloned. This was further confirmed by liquid chromatography mass spectrometry (LCMS) (Appendix 5.1). While it was no longer feasible, due to time constraints, to complete analysis of the original protein sequences, the results gained from the provided plasmids were analysed. The following chapter documents the work done of these proteins with amino acid sequences as shown in Table 5.2.

Plasmid	New Name	pI	MW (Da)	Aliphatic index	Hydropathicity	Instability index
HsPrx3NoTEV (1)	**PRX3A**	6.6	26073	83.4	-0.241	25.7
HsPrx3NoV5 (1)	**PRX3B**	6.5	26065	82.0	-0.218	25.0
HsPrx3NoLink (1)	**PRX3C**	6.5	24623	82.9	-0.194	27.2
HsPrx3Ent (1)	**PRX3D**	6.2	25644	80.0	-0.317	26.3
HsPrx3-8his	**PRX3E**	6.1	26904	85.0	-0.150	26.3
HsPrx3-10his	**PRX3F**	6.1	27174	83.2	-0.176	26.2

Table 5.2: Protein parameters for the constructs expressed by the supplied plasmids. Due to an error in the plasmid manufacturing, these are the proteins that have been expressed. The major difference between the two plasmids is the addition of an N-terminal 6-histdine tag and a thrombin cleavage site. The isoelectric point has increased across all constructs, as has the molecular weight and hydropathicity, with the aliphatic index at a slightly lower value. These are the proteins that have been analysed.

> *PRX3A*

MGSSHHHHHHSSGLVPRGSHMHHHHHHGKPIPNPLLGLDSTAPAVTQHAPY
FKGTAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKANEFHDVN
CEVVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLSDLTKQISRDYGVLLEGSGL
ALRGLFIIDPNGVIKHLSVNDLPVGRSVEETLRLVKAFQYVETHGEVCPANWTPD
SPTIKPSPAASKEYFQKVNQ

> *PRX3B*

MGSSHHHHHHSSGLVPRGSHMHHHHHHENLYFQGIDPFTAPAVTQHAPYFK
GTAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKANEFHDVNCE
VVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLSDLTKQISRDYGVLLEGSGLAL
RGLFIIDPNGVIKHLSVNDLPVGRSVEETLRLVKAFQYVETHGEVCPANWTPDSPT
IKPSPAASKEYFQKVNQ

> *PRX3C*

MGSSHHHHHHSSGLVPRGSHMHHHHHHHAPAVTQHAPYFKGTAVVNGEFKDL
SLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKANEFHDVNCEVVAVSVDSHFSH
LAWINTPRKNGGLGHMNIALLSDLTKQISRDYGVLLEGSGLALRGLFIIDPNGVIK
HLSVNDLPVGRSVEETLRLVKAFQYVETHGEVCPANWTPDSPTIKPSPAASKEYF
QKVNQ

> *PRX3D*

MGSSHHHHHHSSGLVPRGSHMHHHHHHDYKDDDDKAPAVTQHAPYFKGTA
VVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKANEFHDVNCEVVA
VSVDSHFSHLAWINTPRKNGGLGHMNIALLSDLTKQISRDYGVLLEGSGLALRGL
FIIDPNGVIKHLSVNDLPVGRSVEETLRLVKAFQYVETHGEVCPANWTPDSPTIKP
SPAASKEYFQKVNQ

> *PRX3E*

MGSSHHHHHHSSGLVPRGSHMHHGKPIPNPLLGLDSTENLYFQGIDPFTAPA
VTQHAPYFKGTAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSDKA
NEFHDVNCEVVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLSDLTKQISRDYG
VLLEGSGLALRGLFIIDPNGVIKHLSVNDLPVGRSVEETLRLVKAFQYVETHGEVC
PANWTPDSPTIKPSPAASKEYFQKVNQ

> *PRX3F*

MGSSHHHHHHSSGLVPRGSHMHHHHGKPIPNPLLGLDSTENLYFQGIDPFTA
PAVTQHAPYFKGTAVVNGEFKDLSLDDFKGKYLVLFFYPLDFTFVCPTEIVAFSD
KANEFHDVNCEVVAVSVDSHFSHLAWINTPRKNGGLGHMNIALLSDLTKQISRD
YGVLLEGSGLALRGLFIIDPNGVIKHLSVNDLPVGRSVEETLRLVKAFQYVETHGE
VCPANWTPDSPTIKPSPAASKEYFQKVNQ

New construct sequences with an extra N-terminal 6-histdine tag and a thrombin cleavage site

5.3 Histidine content variations

5.3.1 Expression and purification of PRX3E and PRX3F

The genes, cloned into pET 28a(+) between the NdeI and XhoI restriction sites, were synthesised by Epoch Life Science. The open reading frame was cloned into a different expression vector to *Hs*Prx3-6his due to the original vector being discontinued by the company supplying the plasmids.

Host *E. coli* chemically competent cells (BL21 (DE3) ROSETTA, Novagen, 2013; genotype – F- *ompT hsdS*B[B- mB-] *gal dcm* [DE3] pRARE [CamR]) were transformed with the plasmids using heat shock (Chapter 2, at 2.2.2.2). These cells were chosen due to the pRARE codon, which supplied tRNA for the rare codons of *Hs*Prx3 and the presence of T7 RNA polymerase, which allowed expression from the pET vector. The protein was successfully expressed in standard LB media (Chapter 2 at 2.3.1.1) after shaking at 37^{o}C for 4 hours, in the presence of kanamycin and chloramphenicol, and then at 26^{o}C, with the addition of IPTG, for a further 20 hours.

Cells were lysed using sonication (Chapter 2 at 2.3.2.2) and the insoluble fraction was removed by centrifugation (18 000 x g, 30 minutes). The target proteins were isolated from the soluble crude broth using IMAC (fig. 5.1 A and B). SDS-PAGE analysis of the flow through showed that both constructs bound successfully to the IMAC column and were eluted upon increasing imidazole concentration (fig. 5.2 A and B).

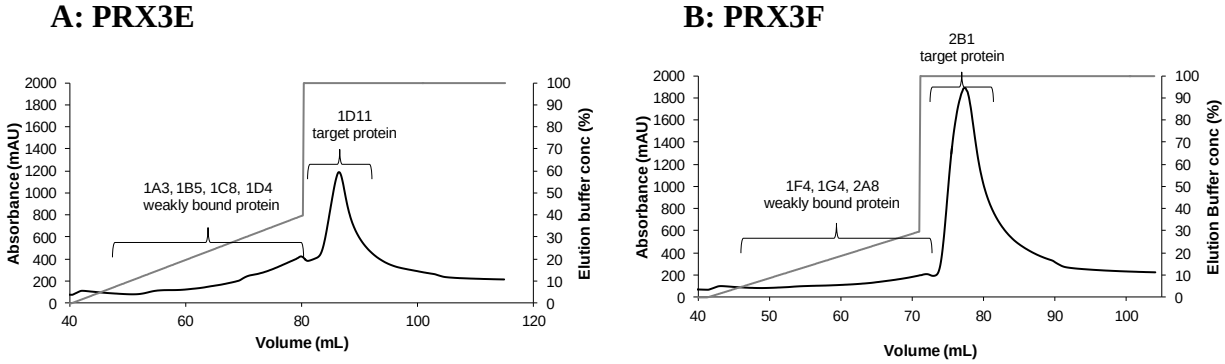


Figure 5.1: Elution of 8-histidine tagged (PRX3E) and 10-histidine (PRX3F) tagged *Hs*Prx3 from IMAC. (A) PRX3E and (B) PRX3F elution from a Ni^{2+} IMAC column upon increasing imidazole concentration. Black: absorbance at 280 nm (mAU); grey: percentage concentration of 0.5 M imidazole buffer (%). Both constructs have been expressed and purified successfully. The presence of target protein is confirmed by SDS-PAGE analysis.

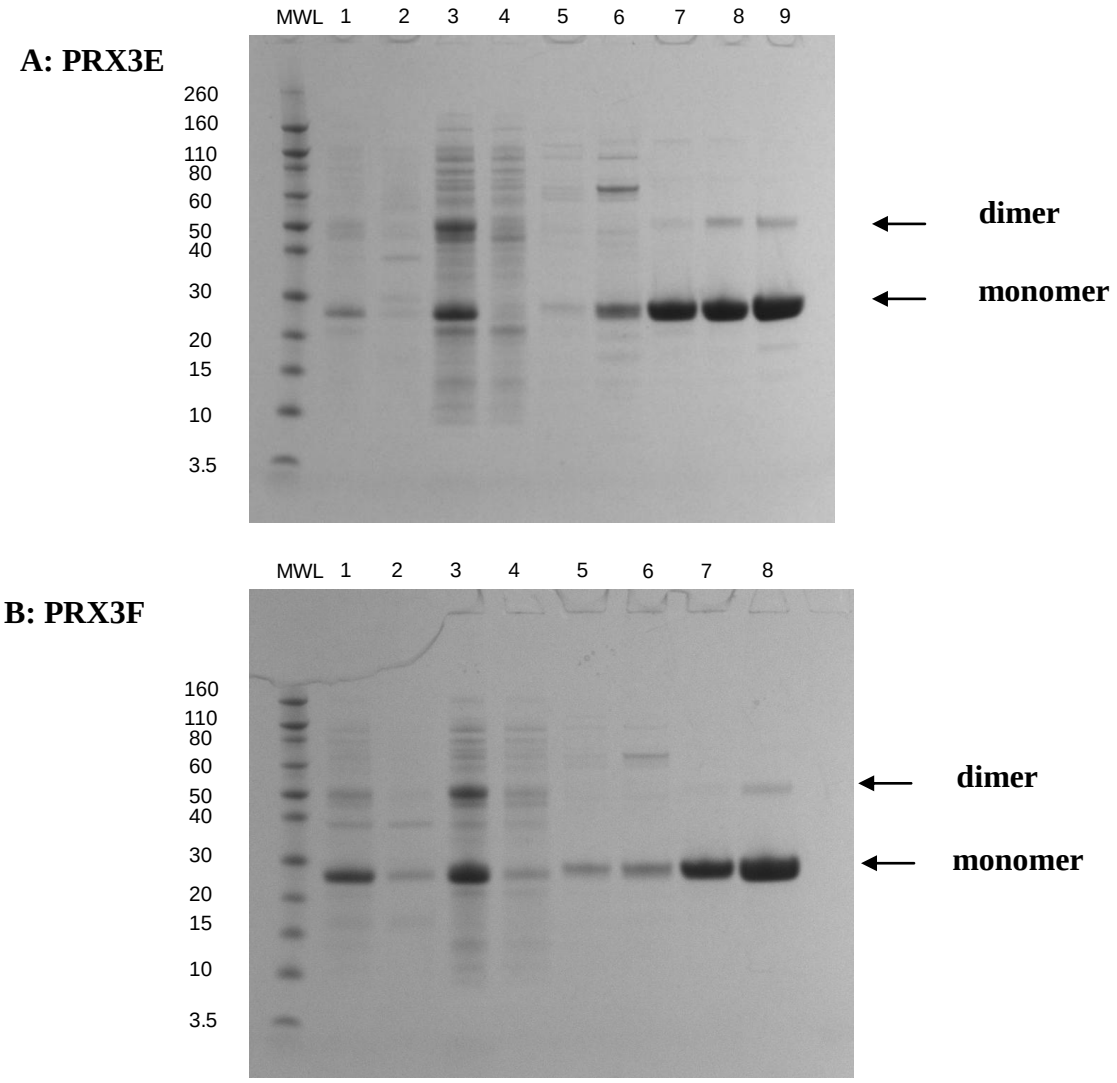


Figure 5.2: SDS-PAGE analysis of 8-histdine tagged (PRX3E) and 10-histidine (PRX3F) tagged *Hs*Prx3 purification. (A) PRX3E purification; 1: crude lysate; 2: insoluble crude; 3: soluble crude; 4: IMAC flow through; 5-8: weakly bound protein; 9: target protein. (B) PRX3F purification; 1: crude lysate; 2: insoluble crude; 3: soluble crude; 4: IMAC flow through; 5-7: weakly bound protein; 8: target protein. Both constructs elute as pure protein upon increasing imidazole concentration. The SDS-PAGE shows the presence of a strong dimer, even under reducing conditions. This could be due to trace amounts of leached nickel from the IMAC column.

For the final purification step, the protein was further isolated by SEC (fig. 5.3 A and B). This was also the desalting step by which the protein was transferred into the storage buffer and the excess imidazole was removed (Chapter 2 at 2.3.2.6). The profile on the SEC trace

displayed a major peak eluting at a volume corresponding to the MW of a single dodecameric species with a shoulder peak, representing a HMW oligomer (fig. 5.3 A and B). There was also a small dimer peak. The elution profile was similar to *Hs*Prx3-6his with the exception of the slightly broader and less well resolved peaks in PRX3F. This indicated a mixture of assemblies in an equilibrium that was faster than the SEC separation time scale. With an increasing number of histidine residues (6his to 10his) within the linker sequence, it is possible that the degree of oligomerisation via metal ion coordination (Ni^{2+} leached from the IMAC) was more pronounced for PRX3F.

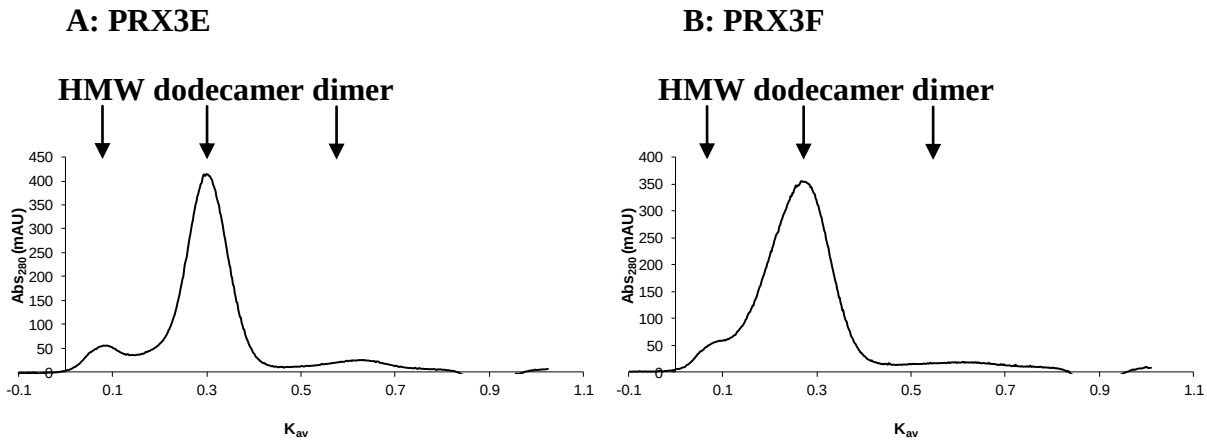


Figure 5.3: SEC trace of 8-histidine (PRX3E) and 10-histdine (PRX3F) tagged *Hs*prx3 eluted IMAC peak. (A) PRX3E and (B) PRX3F. Both peaks are broader than was seen for the 6- histidine construct indicating a mixture of HWM and LWM species. There is a large shoulder peak in the HMW range and a small peak around the molecular weight of a dimer. Removal of trace nickel through EDTA chelation removes the shoulder peak (fig. 5.4 A and B).

5.3.2 Non reduced PRX3E and PRX3F are comparable to *Hs*Prx3-6his at pH 8.0

To ensure that the behaviour of the histidine tag variation constructs was comparable to the 6-histidine tagged form, the characteristics of non-reduced PRX3E and PRX3F at pH 8.0 were evaluated. This was to develop a baseline from which any variations in response to changes in environmental conditions could be measured.

Dodecamer stability and metal chelation conditions

It was important to ensure that changing the tag did not influence the toroid stability under non-reducing conditions, as this was the tecton from which more complex structures formed. In addition to the toroid stability it was important to also assess if the HMW oligomers (seen in the final purification step) were still able to dissociate with the same concentration of chelating agent. The protein was dialysed in pH 8.0 non-reducing buffer to remove the EDTA and any Ni^{2+} leached from IMAC purification. The samples were incubated at $4^{o}C$ overnight to allow for ring disassociation as per the methods in the literature. [24] The change in the histidine tag length did not diminish the stability of dodecamer at pH 8.0 under non-reducing conditions. Furthermore, both the 8-histidine and 10-histidine tagged proteins eluted as a single low molecular weight (LMW) species on SEC. This not only confirmed that the HMW oligomers noted in the final purification step were the result of metal ion coordination, but also showed that the chelating conditions (10 mM ETDA) were sufficient enough to disrupt stacking (fig. 5.4 A and B).

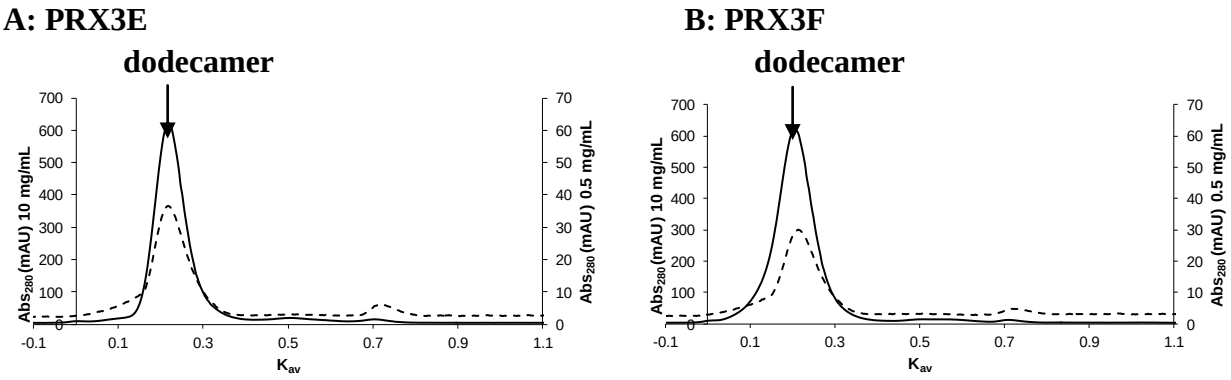


Figure 5.4: SEC trace of 8-histidine (PRX3E) and 10-histidine (PRX3F) tagged *Hs*prx3 at varying concentrations. (A) PRX3E and (B) PRX3F are both in pH 8.0, non-reducing conditions with the EDTA and nickel dialysed out. Black: 10 mg/mL and dotted: 0.5 mg/mL. The constructs elute around the molecular weight of a single LMW ring at both concentrations. Removal of Ni^{2+} using EDTA as the chelating agent has removed the prominent shoulder peak, even at high concentrations.

Protein concentration

The tubes formed from *Hs*Prx3-6his at pH 7.4 and below were sensitive to changes in the protein concentration. With the additional histidine residues in the new sequences it was interesting to investigate if high concentration aided assembly. The protein sample was concentrated to 10 mg/mL, an excess of protein in comparison to *Hs*Prx3-6his analysis, and analysed using SEC. Both proteins eluted as a single peak on SEC with a similar elution profile as seen at 0.5 mg/mL (fig. 5.4 A and B). The size at 10 mg/mL for PRX3F was higher than expected for a single dodecamer and the peak was broad. It is possible that there was a small amount of HMW oligomer present under these conditions. There was a slight shift to a lower MW between 10 mg/mL and 0.5 mg/mL for PRX3F. At low concentrations (0.5 mg/mL), both proteins were behaving in much the same way as *Hs*Prx3-6his at baseline conditions (pH 8.0, 4^{o}C, non-reduced).

Transmission electron microscopy

TEM was used to analyse the quaternary structure of the protein. For both constructs only single rings were seen (fig. 5.5 A and B).

A: PRX3E **B: PRX3F**

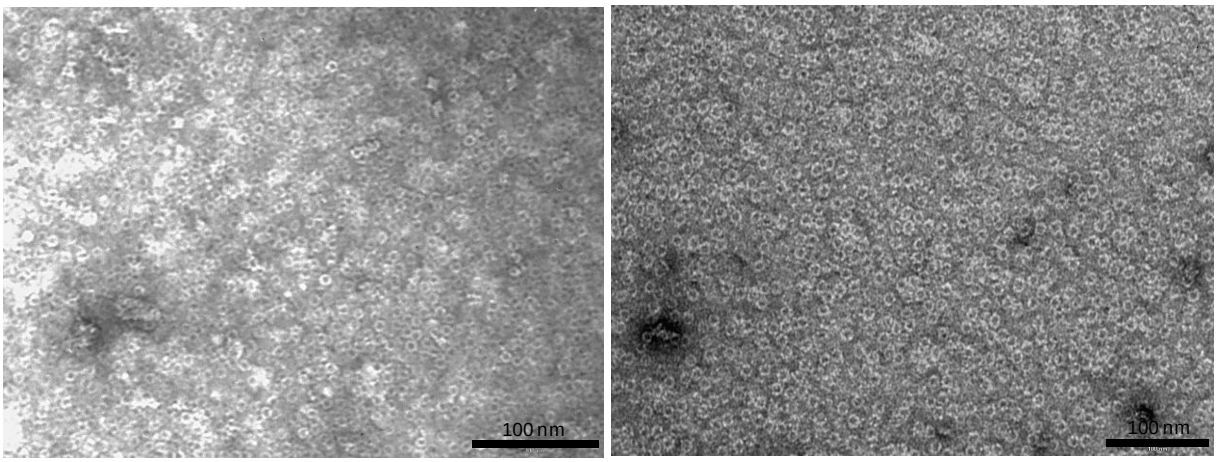


Figure 5.5: TEM of 8-histidine (PRX3E) and 10-histidine tagged (PRX3F) *Hs*Prx3 at pH 8.0 non-reducing conditions. (A) PRX3E and (B) PRX3F were diluted to 100 µg/mL in dH_2O prior to TEM grid preparation and stained with uranyl acetate (1 %). Only single rings are seen. Black line represents 100 nm.

Temperature stability

Finally, all three constructs were analysed using DSF to measure any potential variations in temperature stability. The melting temperature was determined to be 63^{o}C for all three constructs, showing that the change in histidine tag length, plus the additional tag and thrombin cleavage site, did not influence the protein's temperature stability (Appendix 5.2 A, B and C).

All the results confirmed that at baseline conditions the different histidine tag length constructs were comparable in characteristics and behaviour to *Hs*Prx3-6his. The change in plasmid vector did not hinder the recombinant protein stability, and the length of the tag led to no significant changes at pH 8.0 non-reducing conditions.

5.3.3 Increasing the histidine content leads to random aggregation with divalent metal ions

Due to the change in the number of imidazole groups available for coordination it was hypothesised that the degree of oligomerisation would increase. With the different geometry afforded by the change in the tag length there was also the possibility of more inter-toroid tag coordination.

The samples were transferred into pH 8.0 buffer containing 20 mM imidazole and Ni^{2+} was spiked in at a ratio of 1:1 Ni^{2+}: PRX3E or PRX3F (~40 μM). To allow for Ni^{2+} to histidine tag association the mixtures were incubated over night at 4^{o}C before the imidazole was removed via diafiltration. To assess the level of oligomerisation the size of the protein was analysed using AUC (38, 000 rpm). For both constructs, the metal ion coordination appeared to have a lesser influence on protein assembly, with the degree of oligomerisation decreasing when compared to *Hs*Prx3-6his (fig. 5.6 A and B). However, it is important to note that there was a significant amount protein precipitation during this procedure. During the analysis it was attributed to the mechanics of diafiltration. It is now apparent that coordination of Ni^{2+} to the additional imidazoles at the N-terminus was causing the protein to aggregate. It is likely that the protein that was analysed on AUC was the protein that had not coordinated to Ni^{2+} and therefore had not precipitated. To optimise assembly of these constructs via metal coordination a range of different conditions would need to be trialled. These include the Ni^{2+} concentration, initial concentration of imidazole in the assembly buffer, variations in the

buffer exchange, temperature of assembly, and the assembly time. Unfortunately, it was not possible to complete the experiments within the time constraints of the current study.

Interestingly, TEM showed a similar distribution of rings and stacks as *Hs*Prx3-6his (fig. 5.7 A and B). It is possible that the increase in concentration that occurs during grid preparation improved the degree of oligomerisation.

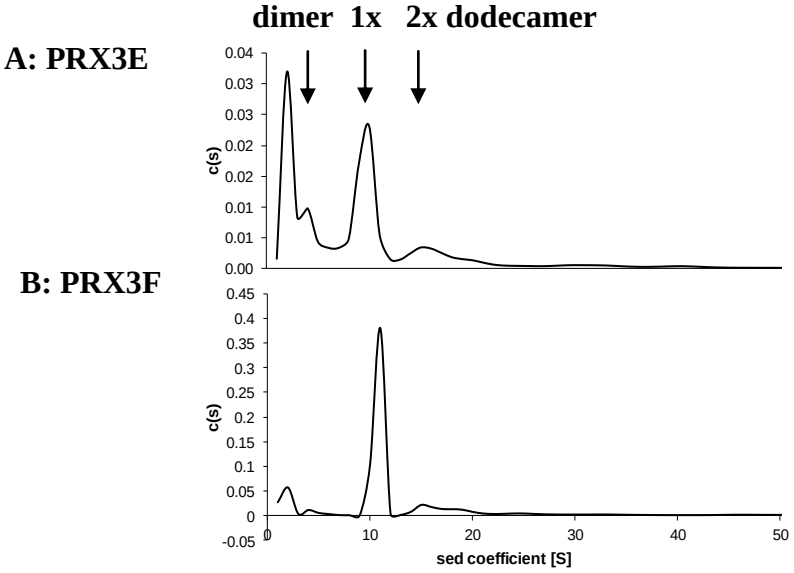
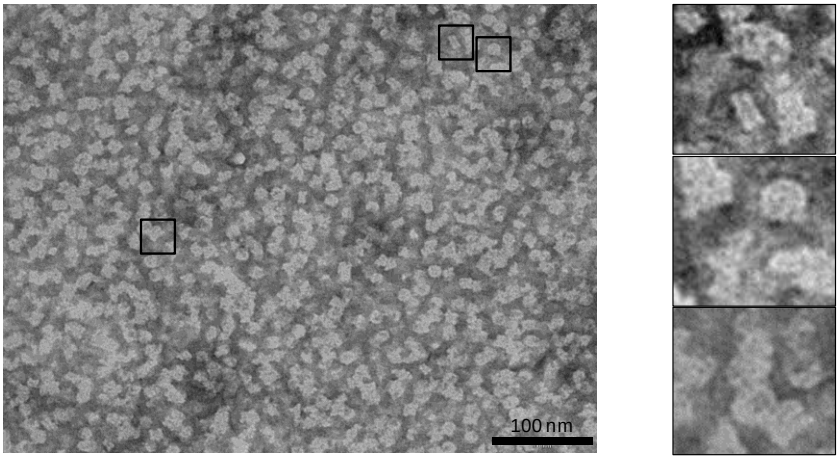


Figure 5.6: AUC trace of 8-histidine (PRX3E) and 10-histidine tagged (PRX3F) *Hs*Prx3 with 1:1 Ni^{2+}: protein monomer (~40 µM). (A) PRX3E and (B) PRX3F. The major species sediments around the molecular weight of a single ring with a small amount of HMW oligomers present. During the removal of imidazole some of the protein crashed out. It is possible that this contained the metal coordinated protein. For PRX3E there is a peak around the molecular weight of a dimer.

A: PRX3E



B: PRX3F

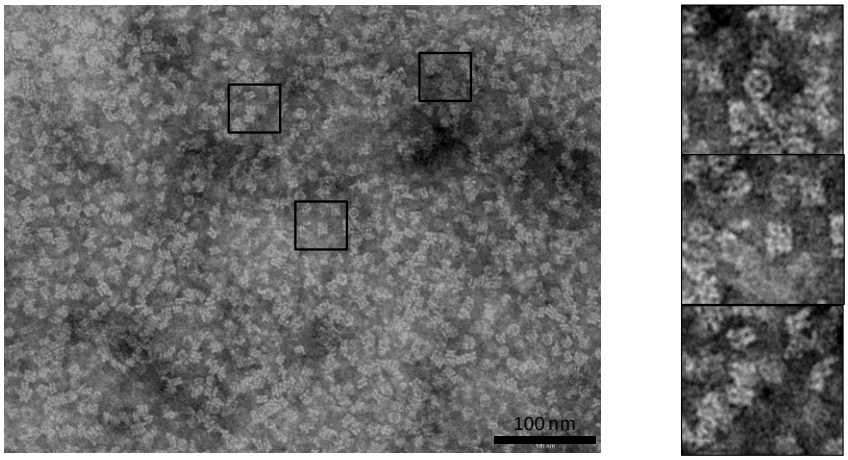


Figure 5.7: TEM of 8-histidine (PRX3E) and 10-histidine tagged (PRX3F) *Hs*Prx3 with 1:1 Ni^{2+}: protein monomer. (A) PRX3E and (B) PRX3F were diluted to 50 μg/mL in dH_2O prior to TEM grid preparation and stained with uranyl acetate (1%). There is a mixture of single rings, some with electron density in the lumen, and double and triple stacks visible. Black line represents 100 nm.

5.3.4 Increasing the histidine content increases the pH sensititivity

The main area of charge residues that differs to the wild type (*Hs*Prx3-WT) construct is the histidine tag. For this reason, it was likely to be involved in some way in the increased pH sensitivity of *Hs*Prx3-6his. To determine if this was the case both constructs were analysed,

initially using SEC at pH 7.4 (Appendix 5.3 A and B) to see if any HMW oligomers were forming, and then at pH 7.4 and 7.2, on AUC and TEM.

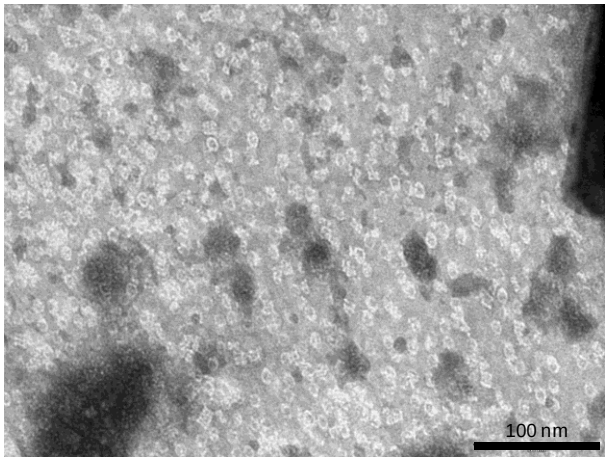
Transmission electron microscopy

PRX3E and PRX3F were transferred into the appropriate buffer with pH values ranging from 7.8 – 7.2. The sample concentration and pH was checked before and after dialysis to ensure complete buffer exchange, and to show that there was no protein precipitation. The samples were diluted to 50 μg/mL in dH_2O directly before TEM grid preparation. The grids were prepared as per the methodology (Chapter 2 at 2.5.4) using 1 % uranyl acetate as the heavy metal stain. TEM images showed a similar spread of different oligomers to *Hs*Prx3-6his at pH 7.8 and 7.6, with single rings present at pH 7.8 and a mixture of rings and stacks of two or three dodecamers at pH 7.6 (fig. 5.8 and 5.9 A and B for PRX3E and PRX3F, respectively). Upon lowering the pH to 7.4 the stacks began to increase in length, and at pH 7.2 elongated 1D tubes had formed (fig. 5.8 and 5.9 C and D for PRX3E and PRX3F, respectively).

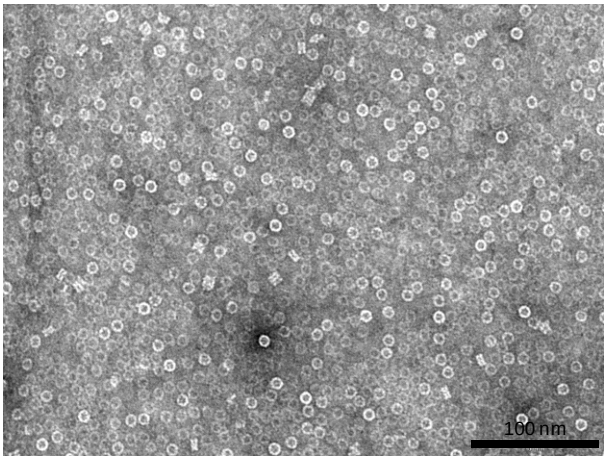
The first major difference between the proteins with varying tag lengths was noted at pH 7.4, with the tube lengths appearing to elongate when the number of histidine residues in the N-terminal tag was increased.

To confirm this was a genuine observation, a range of EM images were selected for size analysis. The width of a single ring along the D6 axis was measured and used to quantify the number of rings per tube. Inputting this information onto a histogram showed a clear trend towards increasing tube length at pH 7.4, with increasing histidine content (fig. 5.10 A). There was a further shift in tube length upon decreasing the pH from 7.4 to 7.2 (fig. 10 B).

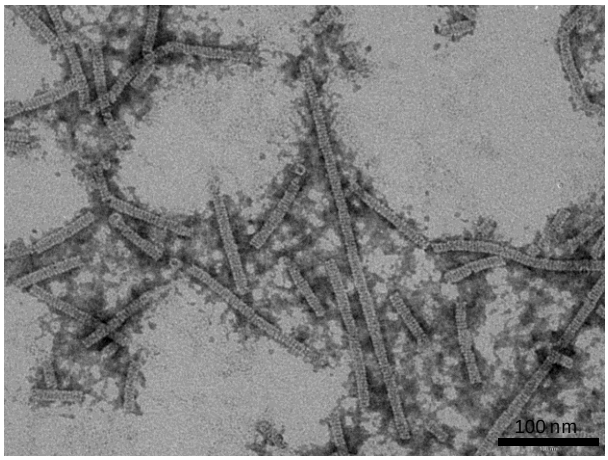
A: pH 7.8



B: pH 7.6



C: pH 7.4



D: pH 7.2

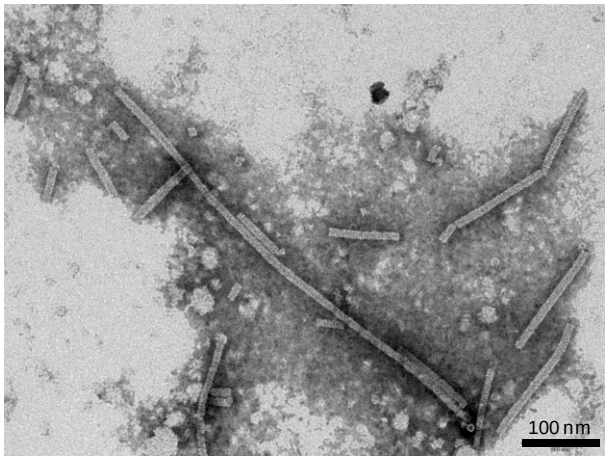
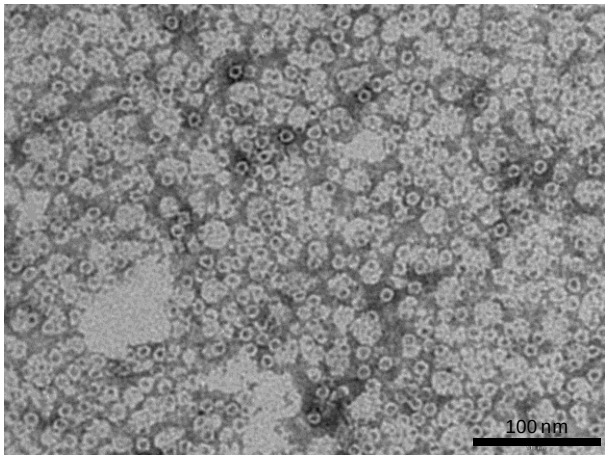
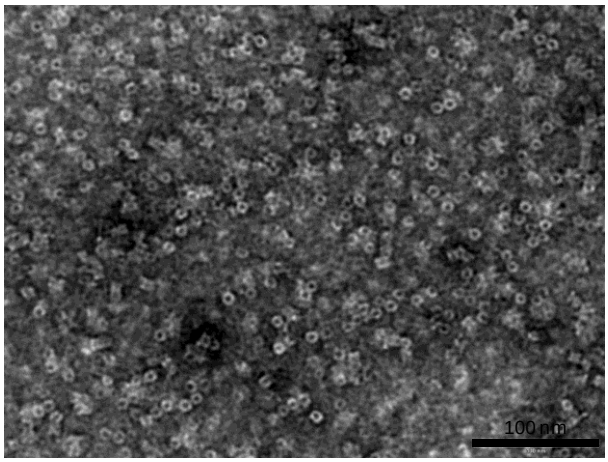


Figure 5.8: TEM of 8-histidine tagged *Hs*Prx3 (PRX3E) at varying pH. (A) pH 7.8, (B) pH 7.6, (C) pH 7.4 and (D) pH 7.2 were diluted to 50 µg/mL in dH_2O prior to TEM grid preparation and stained with uranyl acetate (1%). The tube length increases with increasing pH with the switch between stacks and tubes occurring between pH 7.6 and 7.4. Black line represents 100 nm.

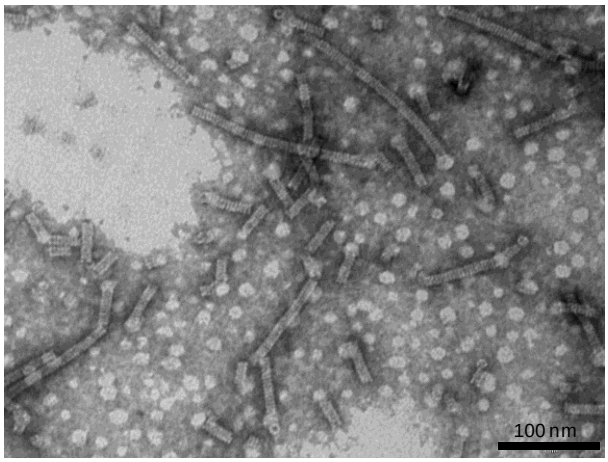
A: pH 7.8



B: pH 7.6



C: pH 7.4



D: pH 7.2

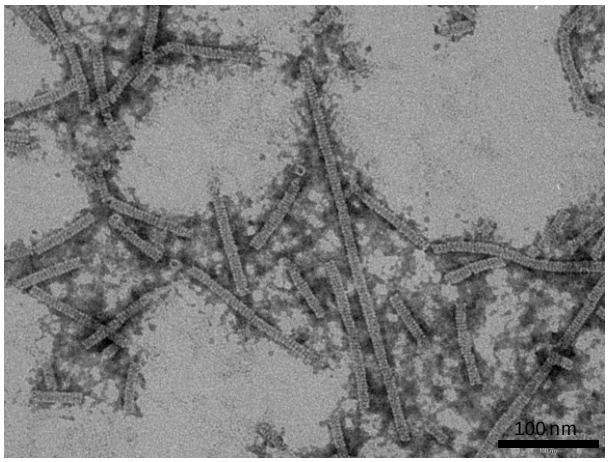


Figure 5.9: TEM of 10-histidine tagged *Hs*Prx3 (PRX3F) at varying pH. (A) pH 7.8, (B) pH 7.6, (C) pH 7.4 and (D) pH 7.2 were diluted to 50 μg/mL in dH_2O prior to TEM grid preparation and stained with uranyl acetate (1%). The tube length elongates with increasing pH, with the switch between stacks and tubes occurring between pH 7.6 and 7.4. Black line represents 100 nm.

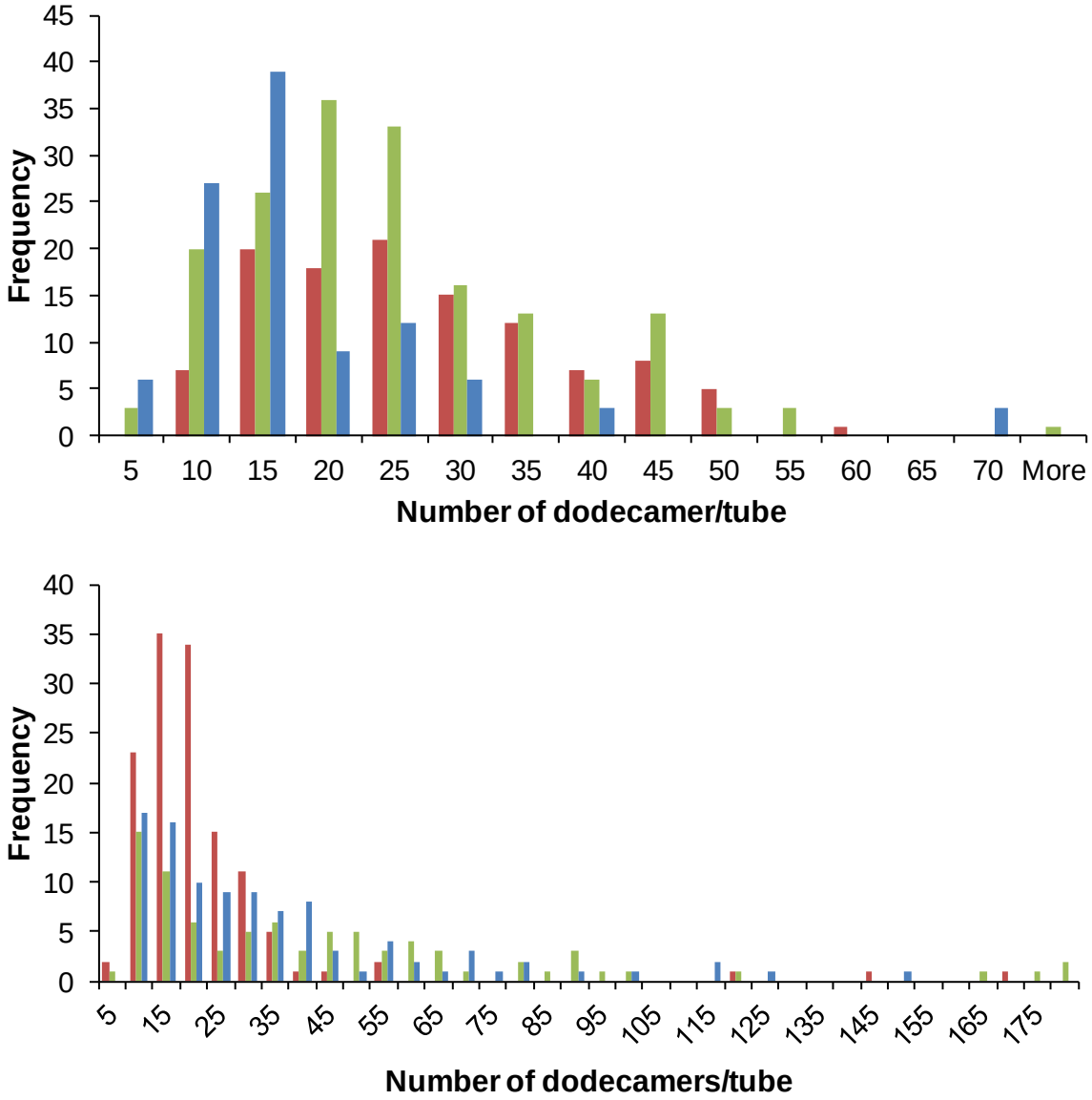


Figure: 5.10: Histograms showing the variation of in extent of tube assembly, as seen on TEM, relative to the histidine tag length; red: 6-histidine; blue: 10-histidine and green: 8-histidine. (A) Tube length at pH 7.4. The tubes are less heterogeneous and shorter in length than those formed at pH 7.2. PRX3E and PRX3F have the longest tubes, validating the trend between tube length and histidine tag length. (B) Tube length at pH 7.2; all tubes are heterogeneous in length with the most prevalent length being between 10 – 20 x dodecamers. The constructs with longer tags are similar in size with PRX3E forming slightly longer tubes. The tubes formed from PRX3E and PRX3F are more heterogeneous than those formed from *Hs*Prx3-6his.

Analytical ultracentrifugation

To ensure the HMW oligomers were stable in solution, the proteins at pH 7.4 and 7.2 were analysed using AUC. Due to the large nature of these assemblies, the centrifuge was set at a lower velocity than the standard measurements (20,000 rpm instead of 38,000 rpm). The samples were prepared as described above, and the MW of the oligomers was approximated from the sedimentation coefficient. The samples were also filtered before AUC analysis to ensure that there was no precipitation. As the traces showed a heterogeneous mixture of species, the regularisation by the maximum entropy and regularisation by the 2nd derivative were utilised to gain a line of best fit.

At pH 7.4 the oligomers for both PRX3E and PRX3F were significantly larger than *Hs*Prx3-6his. The most populated states were within the 5 – 7 dodecamer MW range, with the distribution of the oligomers being similar for both constructs (fig. 5.11 A and B).

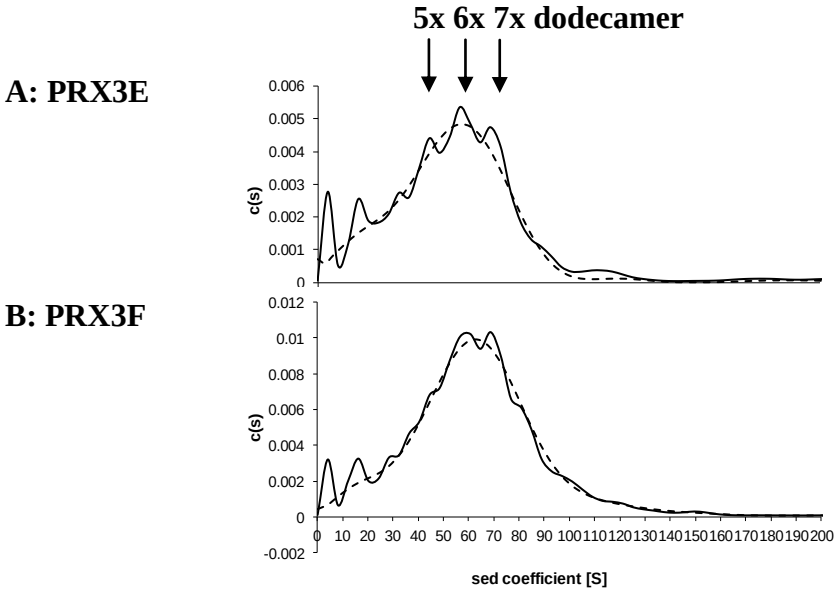


Figure 5.11: AUC of 8-histidine (PRX3E) and 10-histidine tagged (PRX3F) *Hs*Prx3 at pH 7.4. (A) PRX3Eand (B) PRX3F; black: regularisation by the maximum entropy; dotted: regularisation by the 2nd derivative. For both constructs there is only a small amount of LMW oligomer present with the tubes at approximately 5, 6 and 7 x the dodecamer molecular weight being the most prevalent. Regularisation by the 2nd derivative shows that both samples are heterogeneous. The data shows that the large oligomers are stable in solution.

Again, there was a further shift towards HMW oligomers upon decreasing the pH from 7.4 to 7.2 with the most populated states ranging from 8 – 10 dodecamers (fig. 5.12 A and B). The difference between the upper-end of the size ranges was more pronounced for PRX3F, with peaks extending to MWs upwards of 40 x dodecameric stacks, far larger than anything seen for *Hs*Prx3-6his. For both constructs there was a small speak, relating to the MW of a single ring ,but the majority of the dodecamers had stacked into higher ordered structures.

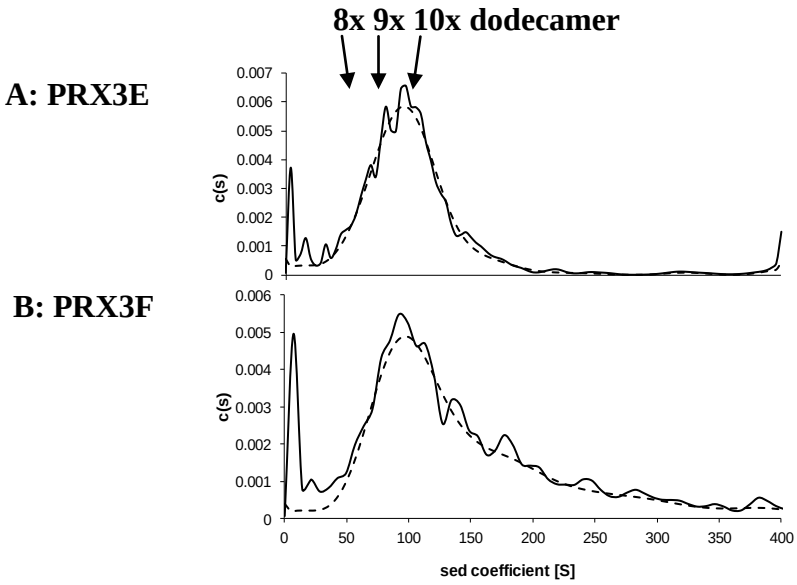


Figure 5.12: AUC of 8-histidine (PRX3E) and 10-histidine tagged (PRX3F) *Hs*Prx3 at pH 7.2. (A) PRX3Eand (B) PRX3F; black: regularisation by the maximum entropy; dotted: regularisation by the 2nd derivative. For both constructs there is only a small amount of LMW oligomer present with the tubes at approximately 8, 9, and 10 x the dodecamer molecular weight being the most prevalent. Regularisation by the 2nd derivative shows that both samples are heterogeneous. The tubes formed at pH 7.2 are longer than those formed at 7.4.

While there is some variability between the different experiments due to the heterogeneous mixture of structures seen, the differences between the protein HMW oligomers were deemed greater than would be expected for mere batch-to-batch variability. Overall it would appear that by increasing the number of histidine residues, and therefore the number of charged

residues in the N-terminal tag, this led to an increase in the size of the structures formed. This agrees well with the hypothesis that the pH sensitivity is originating from the non-native histidine residues within the tag.

5.3.5 Increasing the number of histidine residues increases the rate of assembly

To try and understand the reasoning behind this trend a series of experiments were run whereby the tube growth was measure over time on DLS.

Because the starting point had to be at $t_{0,}$ the usual method of exchanging buffer through dialysis could not be used. This meant that controls needed to be run to assess the differences between the methods. Buffer at pH 7.2 was spiked into a pH 8.0 blank buffer to determine the ratio that was needed for a complete pH switch. The pH of the protein samples was adjusted according to this ratio, and the pH of the sample was checked before further analysis. So as to study one variable at a time, the samples were incubated at 4^{o}C overnight while another batch of protein was simultaneously dialysed into the appropriate buffer. At this time the hydrodynamic radius (D_H) of the protein was measured to check for comparability between these two methods.

	pH 7.2 O/N (20^{o}C)		**pH 7.2 O/N (4^{o}C)**		**pH 7.2 dialysed (4^{o}C)**	
	Z-Ave (Pk 1 mean)	PdI	Pk 1 mean	PdI	Z-Ave	PdI
***Hs*Prx3- 6his**	34.78 ± 0.35	0.296	29.64 ± 1.62	0.555	30.88 ± 0.26	0.114
PRX3F	32.49 ± 2.62	0.690	38.49 ± 9.83	1.000	47.36± 22.26	0.700
PRX3E	53.02 ± 0.51	0.250	74.95 ± 39.53	0.717	85.8 ± 0.89	0.295

Table 5.3: DLS measurements showing the difference between methods of inducing a pH switch. The *samples with a PdI < than 0.3, the hydrodynamic radius for peak 1 was measured instead of the Z-average.*

For *Hs*Prx3-6his, both the methods of pH change, overnight dialysis or dilution, gave comparable values on DLS (Z-ave ~30) (Table 5.3). Interestingly, PRX3E and PRX3F displayed variability in sizes between the different methods, with the sizes of tubes from the dilution method being lower than those seen for buffer exchange via overnight dialysis. It is possible that the slow exchange in buffer facilitated a more controlled assembly, which increased the stability of the structures formed. Due to the measurements on DLS being made at 20^{o}C, the samples were exchanged into a different pH by dilution and left over night at this temperature as an additional control. For the 6-histidine construct this lead to a slight increase in the hydrodynamic radius. For the other proteins it appeared to have a detrimental effect, with the size decreasing (Table 5.3). There is a degree of variability in the assembly of these heterogeneous tubes, and therefore the differences in D_H could also be attributed to batch-to-batch deviation.

Due to the variations in the assembly in response to changes in the buffer exchange protocol it was not possible to directly compare the kinetics of assembly to the standard overnight dialysis method of assembly. However, it was still possible to measure changes in the assembly rate by testing the histidine residue content as the only variable.

The increase in hydrodynamic radius at pH 7.2 was measured over the course of 4 hours to note the differences in behaviour of the proteins. The first major variation was an initial jump in size upon dilution into the pH 7.2 buffer, which was significantly greater for the longer histidine tagged constructs. The sizes continued to increase over the period of 4 hours, with the PRX3F reaching the highest MW (fig. 5.13).

The concentration that was used for the DLS experiments (0.1 mg/mL) was lower than that used in other solution experiments. Samples with a higher concentration gave results that did not meet the equipment quality criteria, possibly as a consequence of the large particle size which resulted in over-scattering, and therefore a lower concentration was used. Due to the nature of assembly, using a lower protein concentration leads to smaller sized oligomers. As a result the size of assemblies on the DLS are likely to be inherently lower than seen in other experiments. That said, the results showed a clear trend in the sizes that correlated well with all the other analysis, and gave good insight into where these changes were originating.

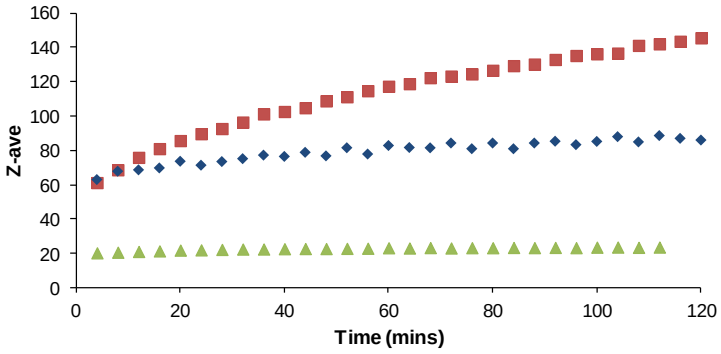


Figure 5.13: Rate of increase in Z-average over time. Green: *Hs*Prx3-6his, blue: PRX3E and red: PRX3F. All samples were diluted into pH 7.2 conditions and the increase in the Z-average was measured over the course of two hours. There appears to be a trend between length of the tag and the size of the oligomers with PRX3F continuing to assemble up to 150 d.nm.

From these results it is possible to design methods towards assembly depending on the types of structures that are needed. If heterogeneity is not an issue, then pH exchange to 7.2 via dilution of the longer tagged constructs would be the best route. The tubes form almost instantaneously but there is a large degree of variability between the sizes. If more homogenous structures are required, then slow pH exchange via overnight dialysis to 7.4 of *Hs*Prx3-6his appears to give the most regular sizes according to TEM and AUC.

5.3.6 Discussion

When designing systems in which proteins assemble in a controllable manner it is imperative to have a good understanding of the correct conditions needed to instigate oligomerisation. The results described here show that the longer histidine tag length constructs form larger structures with a more heterogeneous range of sizes. This correlated well with the hypothesis that the electrostatic interactions at the R-type interface were originating from the histidine residues. As the points of interactions increased, the stability of the interface increased and the structures formed more rapidly.

Without knowing the exact pK_a of these residues, or their exact positioning within the protein structure, it is not possible to assign the precise mechanism of assembly (specifically, the

placement of the tag in respect to the native sequence and the charges that the residues carry). However, this study is not the first instance of histidine tags affecting the folding and oligomerisation of native protein structure. [10, 25] In fact tags, have been implemented in the reduction of enzyme efficiency, [26] the change of the physico-chemical properties of recombinant proteins [10] and the protein dynamics [12], and even reducing the substrate binding efficiency. [11] In a study by Hlima *et al.* it was found that placing a histidine tag at the hydrophobic dimerisation interface of Streptomyces sp SK glucose isomerise destabilised the native interaction, meaning that the dimer did not form. [10] One hypothesis is that 2-Cys peroxiredoxins are predisposed to stack via the R-type interface and that variations, for example structural changes due to overoxidation, are actually stabilising this interaction rather than solely driving assembly. [27] An inverse of the affect seen by Hlima might be at play here. Placing charged residues at an electrostatic interface (R-type interface) [15, 28] may in fact stabilise the interactions that are utilised in the formation of stacks and tubes. With this in mind, it is possible to predict a mechanism of assembly (fig. 5.14):

1. At pH 8.0 the histidine residues carry a net neutral charge. While the toroids intermittently interact, the quaternary structure is not stable and the rate of dissociation is too rapid for the formation of stacks and tubes.
2. Lowering the pH to 7.6 protonates some, but not all of the histidines within the tag. It has been shown that the pK_a of the histidine side chains in a standard poly-6-histidine tag have varying pK_a depending on their proximity to the carbonyl or amine. [20] The charged residues lend some stability to the R-type interface and stacks begin to appear in solution and on EM.
3. At pH 7.4 a high percentage of the histidine residues are protonated, and the high area of positive charge can interact with either negatively charged residues or back-bone carbonyl groups on the adjacent toroid. This interaction is stable in solution, and large structures are seen.
4. At pH 7.2, the majority of the histidine residues within the tags carry a positive charge, stabilising the R-type interface and allowing the formation of elongated 1D tubes.

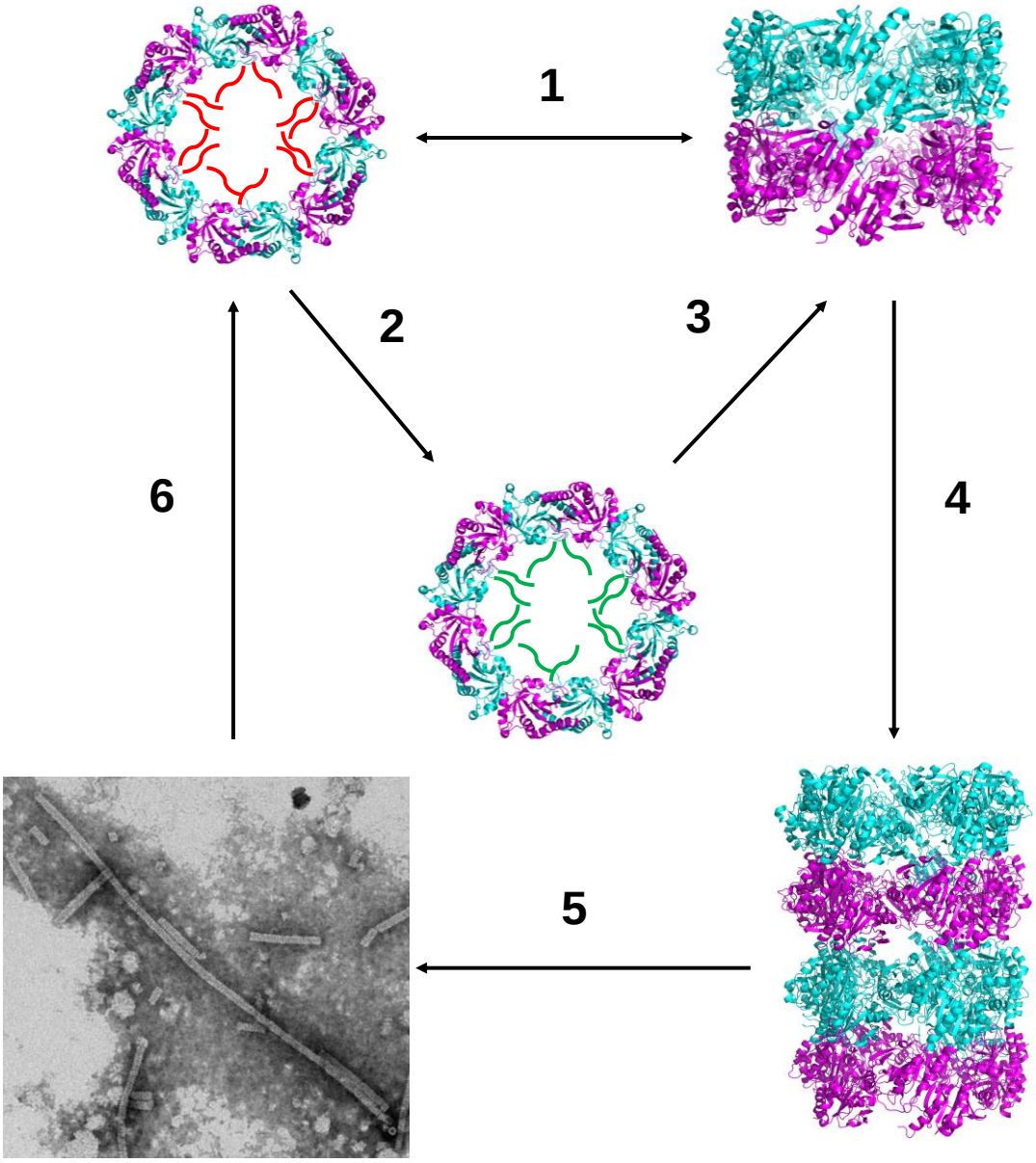


Figure 5.14: Schematic for the possible route of assembly through varying pH. 1 – The protein at pH 8.0 and 7.8. Stacks rapidly assemble and disassemble with interactions at the interface too weak to form stable oligomers. Histidine tags are coloured in red to show the non-charged "off" state. 2 – Reducing the pH causes the histidine within the tags to become charge (green – "on"). 3 - The charge steadily increases from 7.6 – 7.2 leading to increased stabilisation of the R-type interface. 4 – The stacks begin to assemble into large structures at pH 7.4 and 7.2. 5 – Long 1D tubes form at pH 7.2 and high concentration seen on TEM. 6 – The electrostatic interactions are disrupted by high salt concentration and the tubes disassemble back into rings.

5.4 Linker variations

Changes in the linker length and amino acid content were studied to assess how this affected the oligomeric state of the protein. While the tag has been shown to the influence the oligomeric state of *Hs*Prx3's close bovine counterpart (*Bt*Prx3), there are significant differences between the two proteins. The protein quaternary structure in this study is highly sensitive to variations in pH. This was not seen for the bovine construct. The length of the linker is a major point of difference between *Bt*Prx3 and *Hs*Prx-6his, and this was therefore an ideal region to study. With the knowledge that the plasmids were incorrect it is important to note that in addition to variations in linker length, the histidine content was also doubled. The following results will discuss how this affected the quaternary structure and stability of these constructs in pH 8.0 non-reducing conditions.

5.4.1 Expression and purification

The genes, cloned into pET 28a(+) between the NdeI and XhoI restriction sites, were synthesised by Epoch Life Science. Host *E. coli* chemically competent cells (BL21 (DE3) ROSETTA) were transformed with the plasmids using heat shock (Chapter 2 at 2.2.2.2). The proteins were successfully expressed in standard LB media (Chapter 2 at 2.3.1.1) after shaking at 37^{o}C for 4 hours in the presence of kanamycin and chloramphenicol, and then at 26^{o}C with the addition of IPTG, for a further 20 hours.

The proteins were purified using the methods in section 5.3 (Appendix 5.4 A - D) and impurities brought through from the IMAC column were removed by SEC (fig. 5.15 A - C). This was also the desalting step by which the protein was transferred into the storage buffer and the excess imidazole was removed (Chapter 2 at 2.3.2.6). All constructs eluted as a mixture of HMW and LMW oligomers as well as a small amount of dimer. The elution volume of the LMW peak for both PRX3A and PRX3B calculated as a MW that was higher than a single dodecamer indicating a mixture of species. PRX3A and PRX3B also had the highest ratio of HMW to LMW oligomers, with the other two constructs having a higher proportion of dodecamers. The presence of purified target protein in both the HMW and LMW peak was confirmed by SDS-PAGE, with the predominant band being a monomer with a small amount of dimer present (fig. 5.16 A – D). The peak corresponding to the dimer MW

also contained the purified protein. However, for all four constructs there were additional impurities within these fractions and no further work was completed on them.

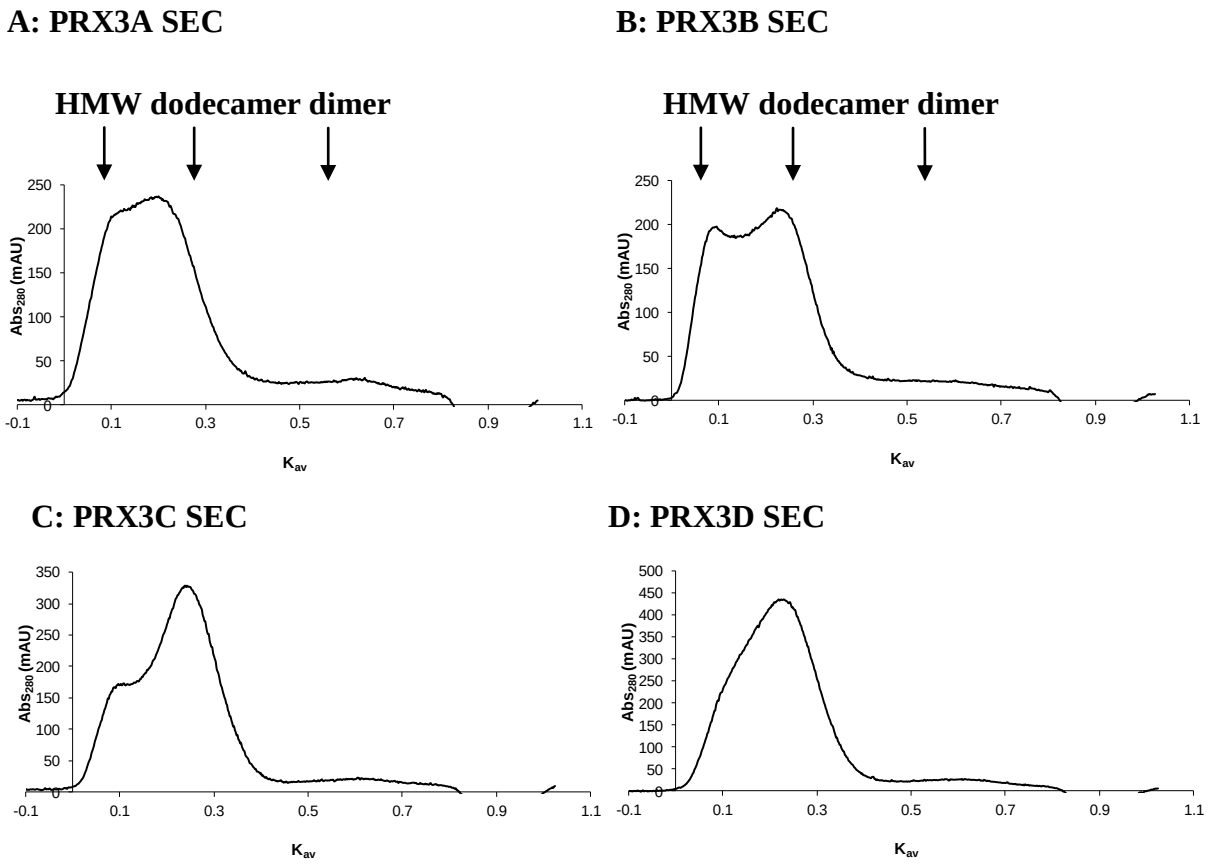


Figure 5.15: SEC chromatography of eluted IMAC peak. (A) PRX3A, (B) PRX3B, (C) PRX3C and (D) PRX3D. All constructs elute as a mixture of HMW and LMW oligomers as well as a small amount of dimer. The apex of the LMW peak for PRX3A and PRX3B is slightly higher in MW than a dodecamer, indicating a mixture of species. PRX3A and PRX3B appear to have the highest ratio of HMW to LMW oligomers, with the other two constructs having a high proportion of dodecamers.

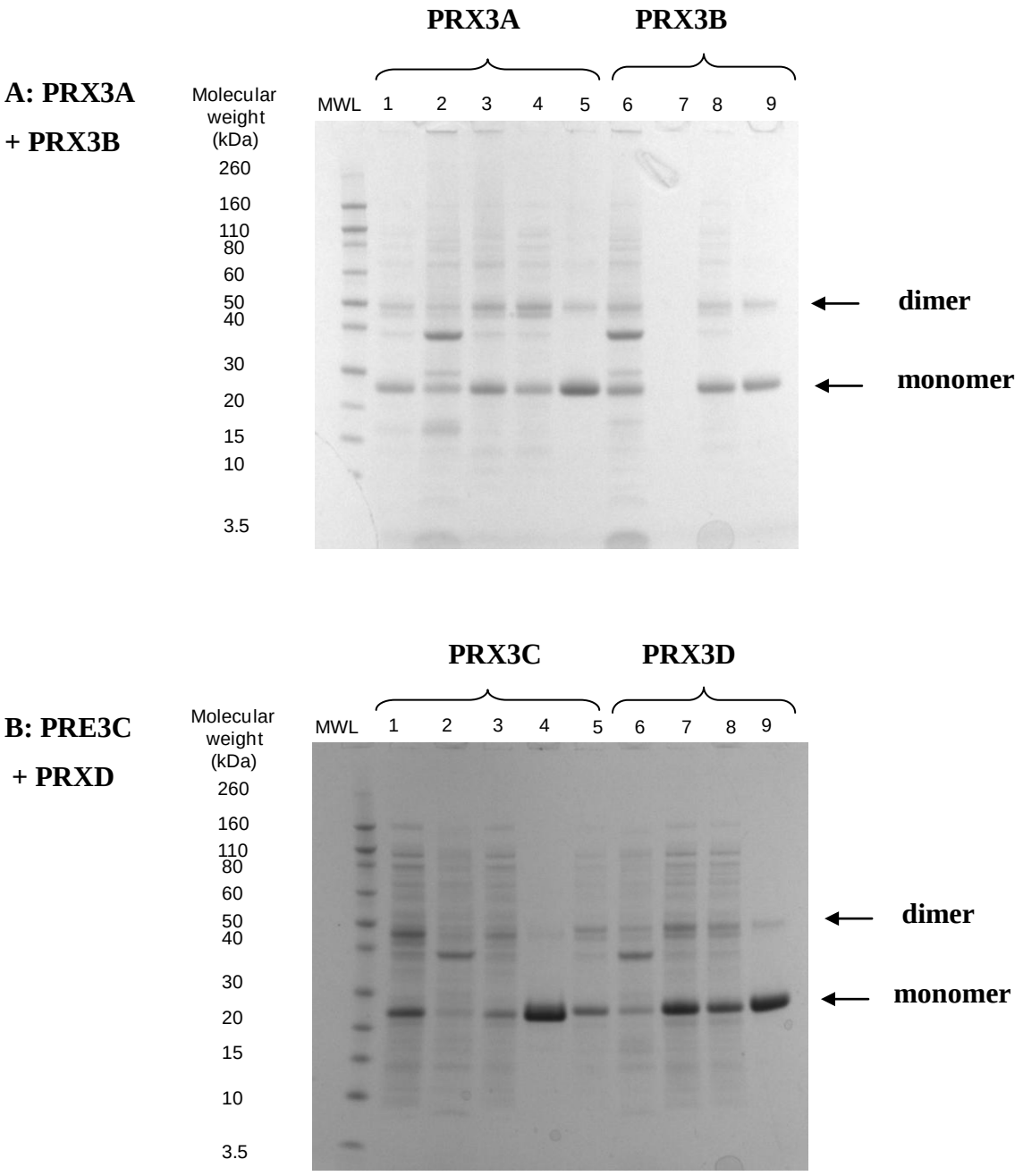


Figure 5.16: SDS-PAGE gel analysis of linker variant purification. (A) PRX3A and PRX3B purification; 1-4: PRX3A crude fractions; 5: isolated PRX3A from IMAC; 6-8: PRX3B crude fractions; 9: isolated PRX3B from IMAC. (B) PRX3C and PRX3D purification; 1-4: PRX3C crude fractions; 5: isolated PRX3C from IMAC; 6-8: PRX3D crude fractions; 9: isolated PRX3D from IMAC. All 4 constructs have been isolated successfully from the IMAC column and run primarily as monomers on denatured electrophoresis, with a small amount of dimer present.

5.4.2 Linker composition influences oligomeric state

The SEC that was completed during the purification process showed that variations in the linker composition and/or the amount of histidine in the tag had a significant effect on the oligomeric state of the protein at pH 8.0.

Protein concentration

HMW species of *Hs*Prx3-6his are sensitive to changes in protein concentration, with larger oligomers forming with increasing concentration. To see if the HMW species seen for the linker variations were also sensitive to variations in protein concentration the new constructs were concentrated to 5 mg/mL, or diluted to 0.5 mg/mL, and their size was measured. The samples were not analysed at 10 mg/mL as at this point some of the protein began to precipitate. This indicates the formation of disordered aggregates and suggests that the change in linker length and histidine residue content had a detrimental influence on the protein stability in aqueous solution at this concentration.

PRX3A, PRX3B, PRX3C and PRX3D had all eluted as a mixture of LMW and HMW oligomers analytical SEC (fig. 5.17 A – D), with PRX3B and PRX3C having the highest and lowest ratio of HMW oligomers at 5 mg/mL respectively (Table 5.4). It is important to note that the mass for the "LMW" peak for both PRX3A and PRX3B was slightly higher than the predicted mass of a dodecamer. It is likely that there was a mixture of single and double rings within this peak which resulted in the peak eluting at a higher MW. Decreasing the protein concentration to 0.5 mg/mL led to a slight decrease in the amount of HMW oligomers seen (fig. 5.17 A – D, Table 5.5). Furthermore, there was a shift towards lower elution volumes in the LMW peak for PRX3A and PRX3B, indicating a possible reduction in concentration of double stacks. However, for all constructs there were still oligomers eluting around the void volume showing assemblies greater than 660 kDa.

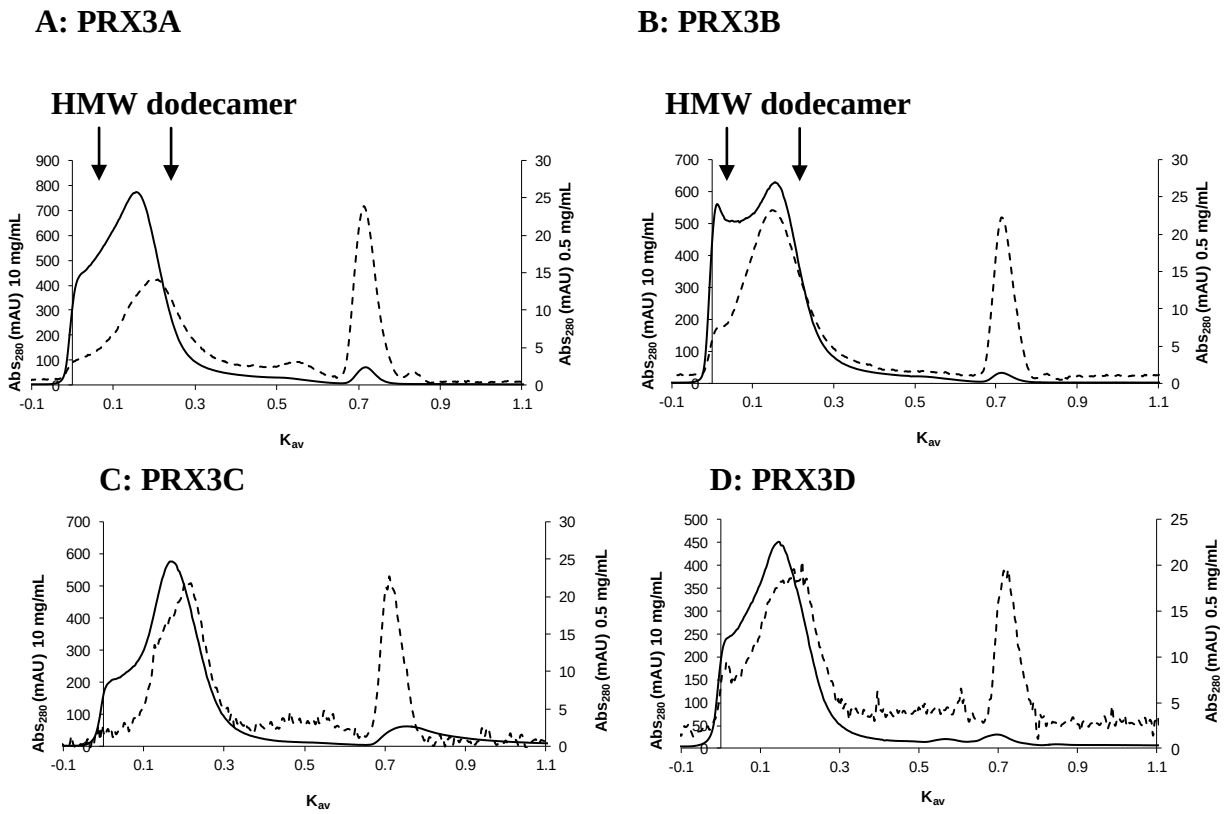


Figure 5.17: SEC traces of linker variation constructs at varying concentrations. (A)PRX3A, (B) PRX3B, (C) PRX3C and (D) PRX3D are all at pH 8.0, non-reducing conditions with the EDTA and nickel dialysed out. Black: 5 mg/mL and dotted: 0.5 mg/mL. For all constructs there is a shift towards the LMW oligomer upon decreasing concentration. However, there is still a large amount of HMW oligomer present, even after the addition of EDTA, indicating that these are formed in a different way to *Hs*Prx3-6his, PRX3E and PRX3F. Integrated values are shown in Table 5.4.

	Area/Peak area (volume) %							
	PRX3A		**PRX3B**		**PRX3C**		**PRX3D**	
	5 mg/mL	**0.5 mg/mL**	**5 mg/mL**	**0.5 mg/mL**	**5 mg/mL**	**0.5 mg/mL**	**5 mg/mL**	**0.5 mg/mL**
HMW	33.02	14.65	35.29	10.53	22.24	14.13	29.56	18.97
LMW	66.98	76.54	64.71	89.47	77.76	85.87	70.44	81.03
Dimer	-	8.81	-	-	-	-	-	-

Table 5.4: Integrated peak values for SEC concentration variations. The volume of the LMW peak (%) increases with decreasing protein concentration.

High salt concentration

It had been shown in the previous chapter (at 4.5.5) that the HMW oligomers of *Hs*Prx3-6his were responsive to changes in salt concentration, likely to be due to the disruption of electrostatic interactions at the R-type interface. To try and assess the origin of the HMW oligomers in response to linker variations, the samples were tested under a range of salt concentrations to determine whether electrostatics were involved in their assembly.

SEC was run on all samples in high salt (600 mM NaCl), non-reduced buffer at pH 8.0. For all the samples there was no significant change in the ratio of HMW and MMW oligomers (Table 5.5, Appendix 5.5) showing that, unlike the stacks and tubes formed from *Hs*Prx3-6his at pH 7.4 and below, electrostatic interactions were not the primary origin of oligomerisation for the linker variation constructs.

	Area/Peak area (volume) %			
	PRX3A	**PRX3B**	**PRX3C**	**PRX3D**
HMW	29.12	31.63	9.02	21.52
LMW	70.88	68.37	90.98	78.48

Table 5.5: Integrated peak values for SEC 600 mM NaCl. All show a mixture of HMW and LMW oligomers. Increasing the salt concentration from 150 mM to 600 mM does not disrupt the interactions that are driving HMW assembly.

AUC

All proteins were further analysed on AUC. This allowed size determination for the oligomers > 660 kDa. The samples were spun at 38,000 rpm and the frictional ratio was set at 1.45, which is the approximate frictional ratio calculated from the crystal structure of *Bt*Prx3 [29] to allow an accurate mass calculation. The size distribution was regularised using both the maximum entropy and 2nd derivative, as described in Chapter 4. Both calculations give a heterogenous mix of discrete species showing that the peaks seen on the trace were real and not artefacts of the Sedfit program (fig. 5.18 A – D).

For all constructs apart from PRX3C, the peak relating to the double stack was the most populated (fig. 5.18 A – D). Surprisingly, there was also a small amount of tetramer seen on all traces that was not seen on SEC. This is possibly due to the dilution of the sample as it is loaded onto a SEC column. [30]

The ratio of each of the different quaternary structures was calculated from the peak integrals (Table 5.6). Stacks of 4 or more dodecamers were calculated as one peak. This is because:

1. There were no discrete peaks.
2. The large oligomers settled within 10 scans at 38,000 x g meaning that there was not enough data within this region to accurately determine MW.

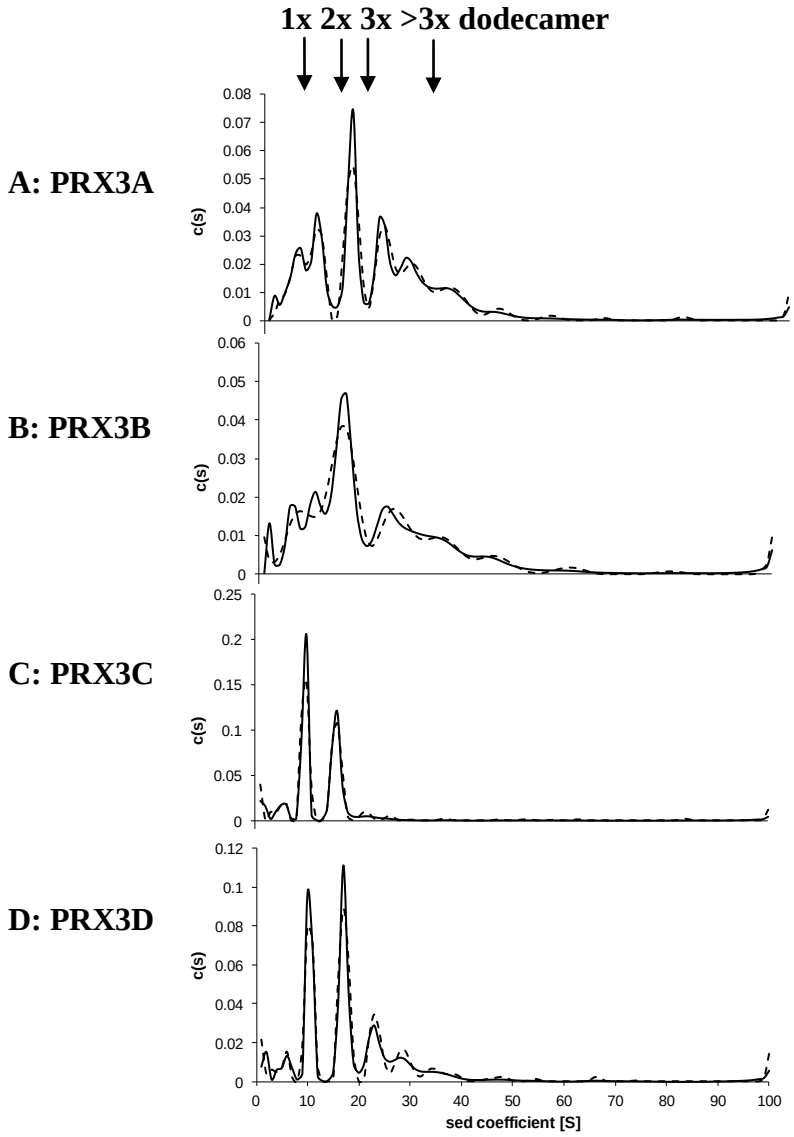


Figure 5.18: AUC of linker variation constructs at 600mM NaCl. High concentration of salt was added to the protein sample (1 mg/mL) in non-reduced buffer pH 8.0. (A) PRX3A, (B) PRX3B, (C) PRX3C and (D) PRX3D all show a mixture of dimer, dodecamer and HMW oligomers. Regularisation by the maximum entropy (black) and by the 2nd derivative (dotted) both show discrete species. PRX3A and PRX3B have the highest proportion of HMW oligomers. A dimer peak is seen for all constructs that was not visible on SEC. Integrated values are shown in Table 5.6.

Construct	tetramer	dodecamer	dodecamer x2	dodecamer x3	> dodecamer x 3
PRX3A	7	14	26	18	35
PRX3B	7	12	29	12	40
PRX3C	7	47	42	4	0
PRX3D	4	26	36	15	19

Table 5.6: Integrate peak areas of PRX3A, PRX3B, PRX3C and PRX3D on AUC. The peak with the largest area is highlighted in yellow

TEM

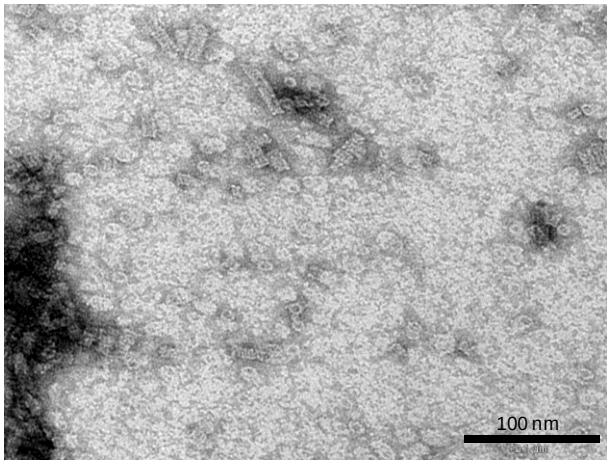
TEM was used to analyse the structure of the HMW and LMW oligomers. The samples were prepared as described in the methodology on Formvar coated copper mesh grids, using 1% uranyl acetate as the negative stain.

For all constructs, a mixture of single rings and stacks were seen. This agreed well with the solution data, with proteins with the most stacks being PRX3A and PRX3B. It was noted that on several of the images there was additional electron density in the centre of the ring (fig.5.19 A-D). This was most prevalent for PRX3B, which had also shown the highest percentage of HMW oligomers on both SEC and AUC (fig. 5.20).

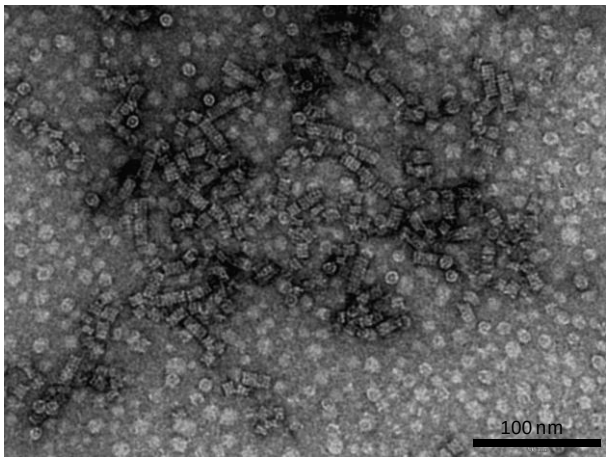
To try and determine whether there was a small amount of impurities in the protein, and therefore an impurity that might be binding the centre of the toroid, a Western blot was run on the proteins. Overloaded SDS-PAGE gel was run in tandem to assess the amount of stable HMW oligomers. An anti-mouse histidine tag antibody was used to identify bands on the SDS-PAGE gel that corresponded to the target protein (Appendix 5.6 A).

The overloaded linker variants all ran as a mixture of monomers, dimers and HMW oligomers. All of these bands were identified as target protein when analysed on Western blot. Additionally, there were small amounts of impurities or protein degradation products seen on SDS-PAGE (Appendix 5.6 B), but the HMW oligomers were stable. This shows that even under harsh denaturing conditions, the proteins still retained some of the quaternary structure. This pertains to a highly stable interface which may explain why the HMW oligomers appear to be so stable and readily formed from the linker variant constructs.

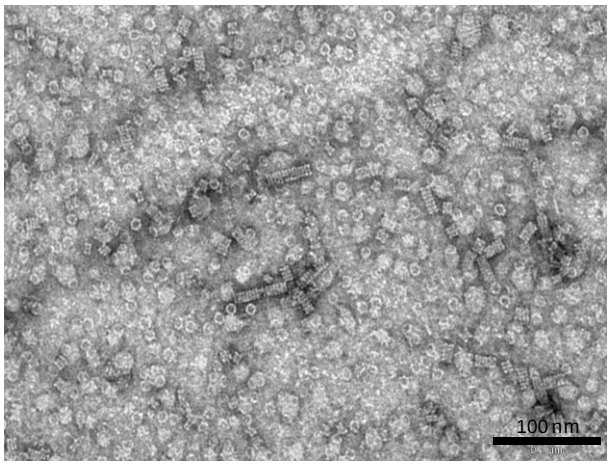
A: PRX3A



B: PRX3B



C: PRX3C



D: PRX3D

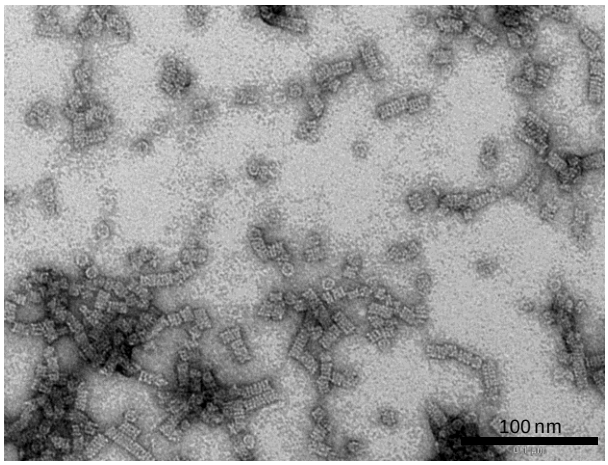


Figure 5.19: TEM of linker variation constructs at pH 8.0 non-reducing conditions. All samples were diluted to 50 μg/mL prior to grid preparation and stained with uranyl actetate (1%). (A) PRX3A, (B) PRX3B, (C) PRX3C and (D) PRX3D all show a mix of LMW and HMW oligomers. This agrees well with the solution data. Across all images there is evidence of single toroids with electron density in the centre. These are most prominent for PRX3B. Black line represents 100 nm.

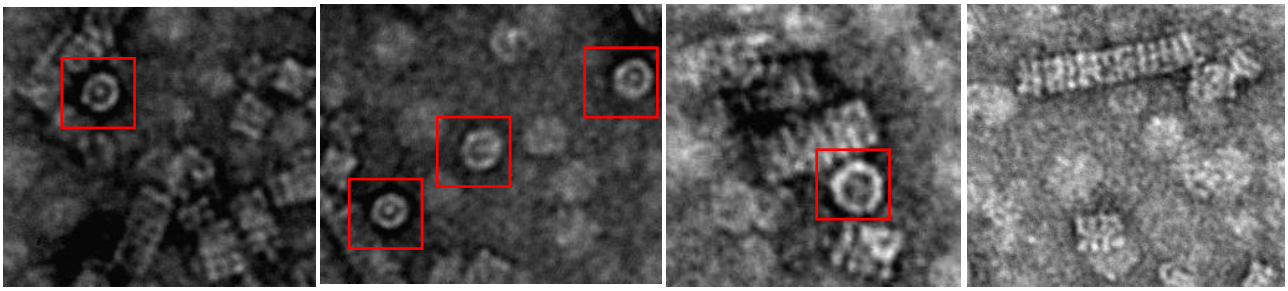


Figure 5.20: TEM of PRX3B. There are a number of different oligomeric states for this construct with a larger number of rings with electron density in the centre (within red squares) as well as stacks and short tubes. This construct had the largest oligomers on AUC as well as TEM. *The images represent the different oligomers formed from PRX3B and are not to scale.*

5.4.3 Discussion

Varying the length of the linker has a marked effect on the protein quaternary structure. This is unlikely to be due solely to the additional histidine tag and thrombin cleavage site, as this was not seen for PRX3E and PRX3F. Furthermore, the addition of high salt did not disrupt the interactions at the R-type interface, and so assembly based exclusively on charge-charge interactions and salt-bridges appeared unlikely. However, it is possible that the increased amount of histidine within the tag led to π-π stacking between the imidazole side groups of adjacent toroids, stabilising the interface. This interaction has been noted at the A-type interface of other peroxiredoxins between neutral histidine residues. [31]

It is also interesting to note that there was some dimer present which was not seen for the full length tag and linker construct. The ring stability gained from the presence of a full length tag and linker may be disrupted when the linker length is varied, further demonstrating that the N-terminal tag can have a long range effect on the overall structure of the protein. [8]

Due to the high degree of electron density within the toroid lumen as visualised of TEM, especially for PRX3B which also had the largest oligomers, it was hypothesised that the presence of a disordered polypeptide chain within the dodecamer lumen may have been instigating a chaperone response. HMW oligomers of *Hs*Prx3 at pH 4.0 have been identified as displaying the characteristics necessary for a putative chaperone. [15] While there is

evidence that LMW peroxiredoxin can act as a holdase chaperone, [18] the majority of reported chaperone species exist as stacks. [17, 19, 28, 32-34] This is believed to be due to the increase in the hydrophobic surface area within stack lumen when compared to a single ring, the binding site for the unfolded or partially unfolded protein substrate. [15, 17, 28, 35] Varying the linker sequence may influence this putative chaperone response and stabilise the peroxiredoxin assembly.

5.5 Structure and function relationship

Over the course of many years of interest in peroxiredoxins, it has been noted that there are a range of different functions that the protein can adopt. [17, 36-39] The structure and oligomeric state is intrinsically linked to the function which is taking place (Chapter 1 at 1.6.6). [40] Of these functions the first to be confirmed, and still considered to be the primary function, is the peroxidase activity. [32, 39, 40]

It has been shown that the switch between dimer and toroid is important for optimal peroxidatic function due to the increased binding efficiency of the thiol reductant to the dimer during the recycling step. [41] This has been further confirmed by the development of mutants with a locked ring structure which show a lower peroxidase activity. [41] The stacking has also been shown to have a detrimental effect on the peroxidatic activity, although the reasoning behind this is not well understood. [33]

One of the benefits of using an enzyme as a proton tecton is that it has in-built function. It is therefore important to discern whether this particular function could be utilised when developing biological nanomaterials.

5.5.1 Peroxidase activity assay

The peroxidase activity of all of the constructs was analysed. This was to determine whether the variations in quaternary structure, linked to the changes in linker and tag length, had a detrimental effect on the enzyme activity. The assay, measured as a competitive kinetics assay with horse radish peroxides (HRP), was developed by Amy Yewdall at the Biomolecular Interaction Centre (BIC) lab at the University of Canterbury, and was based on the Ogusucu method. [42] The colour of HRP changes from orange to clear during the

reduction of H_2O_2 to H_2O and the rate of this reaction is measured by determining the change in absorbance at λ = 403 nm. The samples were prepared as described in the methodology (Chapter 2 at 2.7.6) and each experiment was run in triplicate.

Protein	k_{Prx} (M^{-1} s^{-1})	Standard error (M^{-1} s^{-1})	R^2
***Hs*Prx3-6his**	3.4 x 10^7	1.6 x 10^5	0.95
PRX3A	2.4 x 10^7	9.4 x 10^5	0.97
PRX3B	1.6 x 10^7	9.3 x 10^5	0.94
PRX3C	2.0 x 10^7	7.0 x 10^5	0.98
PRX3D	2.7 x 10^7	1.4 x 10^5	0.96
PRX3E	2.8 x 10^7	1.7 x 10^5	0.95
PRX3F	3.1 x 10^7	1.5 x 10^5	0.96

Table 5.7: The second order rate constants (k_{Prx} M^{-1} s^{-1}) of the peroxidatic activity of all 7 constructs. The competitive activity assay, using HRP as the competitive reactant, was run in accordance to the protocol described in the methodology (Chapter 2 at 2.4.1). k_{Prx} was calculated by plotting $\left(\frac{F}{1-F}\right) k_{HRP}[HRP]$ against [Prx].

All constructs showed fast turn over, with the rate constants ranging from 1.6 x 10^7 to 3.4 x 10^7 $M^{-1}s^{-1}$ (Table 5.7). *Hs*Prx3-6his, PRX3E and PRX3F had the fasted reaction rates, with the rate constants being comparable due to the standard errors. The activity appeared to decrease for the linker length variation constructs. There was also a significantly larger standard error for these constructs. Overall, all enzymes were able to complete their peroxidase activity despite the changes in their quaternary structure that, according to the literature, have the potential to disabled this function. [9, 17, 34]

5.5.2 Oligomerisation in assay buffer conditions

The assay was completed at pH 7.5. With the knowledge that the proteins are sensitive to changes within this pH range, it was important to determine an approximate MW for these samples.

All samples were dialysed overnight into the assay buffer and the MW was analysed by measuring D_H on DLS. As expected, there was a shift towards HMW oligomers at pH 7.5. The Z-ave for a decameric ring at pH 8.0 is ~ 13 d.nm and the smallest oligomer under assay buffer conditions was *Hs*Prx3-6his at 16.47 d.nm. The largest species were PRX3F and PRX3D at 30 and 39 d.nm respectively (Table 5.8). There did not appear to be any correlation between the oligomer size and the activity of the enzyme. PRX3F had one of the highest second order rate constants, and yet was once of the largest oligomers under assay buffer conditions as visualised on DLS. It is important to note that the concentration used on DLS (0.1 mg/mL), is different to the range of concentrations used for activity assay (0.1 mg/mL ~ 4μM). To accurately assess the MW of all of the oligomers in exactly the same conditions used for the assay it would be helpful to analyse the samples using AUC. However, due to the equipment availability it was not possible to complete this within the time frame of this study.

Protein	Z-ave (d.nm)	PdI
***Hs*Prx3-6his**	16.47	0.21
PRX3A	25.27	0.17
PRX3B	26.57	0.23
PRX3C	22.31	0.23
PRX3D	39.91	0.29
PRX3E	23.98	0.17
PRX3F	30.38	0.18

Table 5.8: Size of the protein oligomers in assay buffer conditions. D_H was measure using DLS. All constructs show an increase in size when in assay buffer conditions (pH 7.5, reduced). PRX3D has the highest MW.

5.5.3 Discussion

All seven constructs showed a high peroxidatic activity despite the locked ring and stack conformation. There did not appear to be any correlation between the oligomer size and the activity of the enzyme.

Interestingly, PRX3B showed the lowest activity. This was the construct that also contained a large amount of electron density within the toroid lumen according to TEM. It is possible that this construct, in particular, had more chaperone characteristics. The conditions that are normally used to drive toroid assembly; acidic pH, [24, 28] high concentration of H_2O_2, [17] and significant changes in temperature, [18] may induce structural changes that are detrimental to the peroxidase activity. It is hypothesised that peroxiredoxin stacks are still highly active as long as the active site (and other regions that contribute to enzymatic activity) are in the conformation that allows optimum catalytic activity. [43] With this in mind PRX3B, which appeared to be acting as a chaperone on TEM, [18] would have the lowest enzymatic activity due to a switch in function, and not solely because of an increase in complex quaternary structure. While not within the scope of this project, a crystal structure would be useful to comprehensively analysis this hypothesis. A chaperone activity assay would also be necessary.

The redox activity of the active site could lead to the development of functional materials. For example, the protein could be bound to a surface in a direct or indirect manner utilising the metal chelation tag, or the electrostatic contacts, respectively. This could be used in the development of a high surface area catalyst such as a "green science" catalytic converter, converting H_2O_2 to H_2O. The proton transfer involved in this redox reaction could also be utilised to develop a battery, with the protons carrying the charge.

Overall, the fact that all proteins show high peroxidatic activity provides a wealth of opportunities for developing these tectons into functional nanomaterials.

5.6 Conclusion

While the hypothesis behind the original design of the tag and linker variation constructs could not be tested due to the error in plasmid production, the data gained still gave insight into the mechanism of assembly and the ways in which it could be controlled

With an increase in the amount of histidine available for electrostatic interactions there was an increase in the degree of oligomerisation in response to changes in pH. This further strengthened the idea that stack and tube stabilisation is a consequence of electrostatic interactions between the protonated histidine residues and areas of negative charge on the neighbouring dodecamer. It also provided a means of designing tubes of a certain length within a specified time frame. All of the linker variants stacked at pH 8.0, meaning that studying the assembly at pH lower than this would be difficult. However, it was interesting to note that these changes did have a large impact on the quaternary structure of the protein. For all four constructs there was evidence of SDS stable HMW oligomers that formed stacks and tubes, as visualised by AUC and TEM. Perhaps most intriguingly, there was a significant amount of electron density within the lumen of many of the dodecamers, as visualised by TEM. The tag and linker are likely to exist as disordered peptides. It was postulated that this could be initiating a chaperone response, causing the dodecamers to stack, however further investigation would be needed to confirm this.

There was no discernible correlation between oligomeric state and peroxidase activity. However, the construct that showed the lowest activity, PRX3B, also had the highest amount of toroids with electron density within the lumen. This would suggest that the reduction in peroxidase activity is not directly linked to the quaternary structure, but instead the switch in function of the peroxiredoxin (specifically from a peroxidase to a chaperone).

Overall, it is clear from this study that the presence of a non-native peptide sequence on the N-terminus of *Hs*Prx3 has a large effect on its the oligomeric state. This change in quaternary structure can be harnessed to develop a protein tecton that has a wealth of in-built functionalities.

5.7 References

1. Anfinsen CB: Formation and stabilisation of protein structure. *Biochemical Journal* 1972, 128(4):737-&.
2. Marianayagam NJ, Sunde M, Matthews JM: The power of two: protein dimerization in biology. *Trends in Biochemical Sciences* 2004, 29(11):618-625.
3. Arakawa T, Prestrelski SJ, Kenney WC, Carpenter JF: Factors affecting short-term and long-term stabilities of proteins. *Advanced Drug Delivery Reviews* 1993, 10(1):1-28.
4. Rose GD, Fleming PJ, Banavar JR, Maritan A: A backbone-based theory of protein folding. *Proceedings of the National Academy of Sciences of the United States of America* 2006, 103(45):16623-16633.
5. Cingolani R, Rinaldi R, Maruccio G, Biasco A: Nanotechnology approaches to self-organized bio-molecular devices. *Physica E-Low-Dimensional Systems & Nanostructures* 2002, 13(2-4):1229-1235.
6. Yeates T: Nanobiotechnology: protein arrays made to order. *Nature Nanotechnology* 2011, 6:541 - 542.
7. Gradisar H, Jerala R: Self-assembled bionanostructures: proteins following the lead of DNA nanostructures. *Journal of Nanobiotechnology* 2014, 12(1):4.
8. Gretes MC, Karplus PA: Observed octameric assembly of a Plasmodium yoelii peroxiredoxin can be explained by the replacement of native "ball-and-socket" interacting residues by an affinity tag. *Protein Science* 2013, 22(10):1445-1452.
9. Cao Z, Bhella D, Lindsay JG: Reconstitution of the mitochondrial PrxIII antioxidant defence pathway: General properties and factors affecting PrxII activity and oligomeric state. *Journal of Molecular Biology* 2007, 372(4):1022-1033.
10. Ben Hlima H, Ayadi D, Aghajari N, Bejar S: Differential properties of native and tagged or untagged recombinant glucose isomerases of Streptomyces sp SK and possible implication of the glycosylation. *Journal of Molecular Catalysis B: Enzymatic* 2013, 94:82-87.
11. Majorek KA, Kuhn ML, Chruszcz M, Anderson WF, Minor W: Double trouble-Buffer selection and His-tag presence may be responsible for nonreproducibility of biomedical experiments. *Protein Science* 2014, 23(10):1359-1368.
12. Thielges MC, Chung JK, Axup JY, Fayer MD: Influence of histidine tag attachment on picosecond protein dynamics. *Biochemistry* 2011, 50(25):5799-5805.
13. Watanabe H, Matsumaru H, Ooishi A, Feng YW, Odahara T, Suto K, Honda S: Optimizing pH Response of Affinity between Protein G and IgG Fc. How electrostatic moddulation affect protein-protein interactions. *Journal of Biological Chemistry* 2009, 284(18):12373-12383.
14. Morais MAB, Giuseppe PO, Souza T, Alegria TGP, Oliveira MA, Netto LES, Murakami MT: How pH modulates the dimer-decamer interconversion of 2-Cys Peroxiredoxins from the Prx1 subfamily. *Journal of Biological Chemistry* 2015, 290(13):8582-8590.
15. Radjainia M, Venugopal H, Desfosses A, Phillips AJ, Yewdall NA, Hampton MB, Gerrard JA, Mitra AK: Cryo-electron microscopy structure of human Peroxiredoxin-3 filament reveals the assembly of a putative chaperone. *Structure* 2015, 23(5):912-920.
16. Kumsta C, Jakob U: Redox-regulated chaperones. *Biochemistry* 2009, 48(22):4666-4676.
17. Jang HH, Lee KO, Chi YH, Jung BG, Park SK, Park JH, Lee JR, Lee SS, Moon JC, Yun JW *et al*: Two enzymes in one: Two yeast peroxiredoxins display oxidative

stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell* 2004, 117(5):625-635.

18. Teixeira F, Castro H, Cruz T, Tse E, Koldewey P, Southworth DR, Tomas AM, Jakob U: Mitochondrial peroxiredoxin functions as crucial chaperone reservoir in Leishmania infantum. *Proceedings of the National Academy of Sciences of the United States of America* 2015, 112(7):E616-E624.
19. Konig J, Galliardt H, Jutte P, Schaper S, Dittmann L, Dietz KJ: The conformational bases for the two functionalities of 2-cysteine peroxiredoxins as peroxidase and chaperone. *Journal of Experimental Botany* 2013, 64(11):3483-3497.
20. Svanedal I, Boija S, Almesaker A, Persson G, Andersson F, Hedenstrom E, Bylund D, Norgren M, Edlund H: Metal ion coordination, conditional stability constants, and solution behavior of chelating surfactant metal complexes. *Langmuir* 2014, 30(16):4605-4612.
21. Brodin JD, Ambroggio XI, Tang C, Parent KN, Baker TS, Tezcan FA: Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* 2012, 4(5):375-382.
22. Brodin JD, Smith SJ, Carr JR, Tezcan FA: Designed, helical protein nanotubes with variable diameters from a single building block. *Journal of the American Chemical Society* 2015, 137(33):10468-10471.
23. Huard DJE, Kane KM, Tezcan FA: Re-engineering protein interfaces yields copper-inducible ferritin cage assembly. *Nature Chemical Biology* 2013, 9(3):169-176.
24. Phillips AJ, Littlejohn J, Yewdall NA, Zhu T, Valery C, Pearce FG, Mitra AK, Radjainia M, Gerrard JA: Peroxiredoxin is a versatile self-assembling tecton for protein nanotechnology. *Biomacromolecules* 2014, 15(5):1871-1881.
25. Chant A, Kraemer-Pecore CM, Watkin R, Kneale GG: Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA causes a conformational change at the DNA-binding site. *Protein Expression and Purification* 2005, 39(2):152-159.
26. Chen Z, Li Y, Yuan Q: Study the effect of His-tag on chondroitinase ABC I based on characterization of enzyme. *International Journal of Biological Macromolecules* 2015, 78:96-101.
27. Noichri Y, Palais G, Ruby V, D'Autreaux B, Delaunay-Moisan A, Nystrom T, Molin M, Toledano MB: In vivo parameters influencing 2-Cys Prx oligomerization: The role of enzyme sulfinylation. *Redox Biology* 2015, 6:326-333.
28. Saccoccia F, Di Micco P, Boumis G, Brunori M, Koutris I, Miele AE, Morea V, Sriratana P, Williams DL, Bellelli A *et al*: Moonlighting by different stressors: Crystal structure of the chaperone species of a 2-Cys Peroxiredoxin. *Structure* 2012, 20(3):429-439.
29. Cao Z, McGow DP, Shepherd C, Lindsay JG: Improved catenated structures of Bovine Peroxiredoxin III F190L reveal details of ring-ring interactions and a novel conformational state. *PLOS One* 2015, 10(4).
30. Duong-Ly KC, Gabelli SB: Gel filtration chromatography (size exclusion chromatography) of proteins. In: *Laboratory Methods in Enzymology: Protein, Pt C.* Edited by Lorsch J, vol. 541. San Diego: Elsevier Academic Press Inc; 2014: 105-114.
31. Nguyen JB, Pool CD, Wong CYB, Treger RS, Williams DL, Cappello M, Lea WA, Simeonov A, Vermeire JJ, Modis Y: Peroxiredoxin-1 from the human hookworm Ancylostoma ceylanicum forms a stable oxidized decamer and is covalently inhibited by conoidin A. *Chemistry & Biology* 2013, 20(8):991-1001.

32. Hall A, Karplus PA, Poole LB: Typical 2-Cys peroxiredoxins - structures, mechanisms and functions. *Federation of European Biochemical Societies Journal* 2009, 276(9):2469-2477.
33. Banerjee M, Chakravarty D, Ballal A: Redox-dependent chaperone/peroxidase function of 2-Cys-Prx from the cyanobacterium Anabaena PCC7120: role in oxidative stress tolerance. *BioMed Central plant biology* 2015, 15(1):444-444.
34. Angelucci F, Saccoccia F, Ardini M, Boumis G, Brunori M, Di Leandro L, Ippolitl R, Miele AE, Natoli G, Scotti S *et al*: Switching between the alternative structures and functions of a 2-Cys Peroxiredoxin, by site-directed mutagenesis. *Journal of Molecular Biology* 2013, 425(22):4556-4568.
35. Tavender TJ, Sheppard AM, Bulleid NJ: Peroxiredoxin IV is an endoplasmic reticulum-localized enzyme forming oligomeric complexes in human cells. *Biochemical Journal* 2008, 411:191-199.
36. Wood ZA, Schroder E, Harris JR, Poole LB: Structure, mechanism and regulation of peroxiredoxins. *Trends in Biochemical Sciences* 2003, 28(1):32-40.
37. Karplus PA, Hall A: Structural survey of the peroxiredoxins. *Sub-Cellular Biochemistry* 2007, 44:41-60.
38. Poynton RA, Hampton MB: Peroxiredoxins as biomarkers of oxidative stress. *Biochimica Et Biophysica Acta-General Subjects* 2014, 1840(2):906-912.
39. Karplus PA: A primer on peroxiredoxin biochemistry. *Free Radical Biology and Medicine* 2015, 80:183-190.
40. Perkins A, Poole LB, Karplus PA: Tuning of peroxiredoxin catalysis for various physiological roles. *Biochemistry* 2014, 53(49):7693-7705.
41. Matsumura T, Okamoto K, Iwahara S-I, Hori H, Takahashi Y, Nishino T, Abe Y: Dimer-oligomer interconversion of wild-type and mutant rat 2-Cys peroxiredoxin. *Journal of Biological Chemistry* 2008, 283(1):284-293.
42. Ogusucu R, Rettori D, Munhoz DC, Netto LES, Augusto O: Reactions of yeast thioredoxin peroxidases I and II with hydrogen peroxide and peroxynitrite: Rate constants by competitive kinetics. *Free Radical Biology and Medicine* 2007, 42(3):326-334.
43. Hall A, Parsonage D, Poole LB, Karplus PA: Structural evidence that Peroxiredoxin catalytic power is based on transition-state stabilization. *Journal of Molecular Biology* 2010, 402(1):194-209.

Chapter Six: Conclusions and suggestions for future work

The ability of proteins to associate into 3D functional entities is unparalleled when compared to synthetic analogues. [1-3] Their functional complexity, shape, and the different modes of assembly are all encoded in the amino acid sequence. It is therefore possible to predict systems that will drive controlled association, making proteins attractive tools for the formation of nanomaterials using the bottom-up approach. The aim of this study was to increase the repertoire of available protein building blocks from which these structures could form.

Using proteins that had the ability to assemble in nature meant that the interfaces that guided assembly were already in place. Additionally, selecting proteins that had in-built functions was also seen as beneficial for the formation of functional nanomaterials. The different modes of oligomerisation were examined for three different proteins, Lsr2, Nterm-Lsr2 and *Hs*Prx3. All three proteins displayed features that aligned them with the goals of this work.

Throughout this study, novel and tailorable methods were developed to form protein tectons that associated upon a specific switch. The relatively simple modes of assembly allowed the study of additional routes of exerting control over these systems. Using a range of different environmental factors (protein concentration, pH, ionic strength, and metal chelation) it was possible to develop assembly systems that produced materials of a controlled size. These systems could be utilised in the development of functional nanomaterials in future work.

6.1 Oligomerisation of Lsr2 and its derivatives

Lsr2 was identified as a potential protein tecton due to a number of native modes of assembly that had been identified in the associated literature. [4-7] The ability of this small, basic, histone-like protein to organise DNA into bacterial nucleoids and nucleoprotein fibres also aligned well with the idea of developing protein building blocks with in-built function. [6, 8, 9] During the purification process it was noted that the protein bound to significant amounts of genomic DNA. Due to the fact it was not possible to remove the DNA using the methods

employed in this thesis, the controlled assembly of the N-terminal dimerisation domain was explored instead.

6.1.1 Assembly in response to an enzymatic switch

Nterm-Lsr2 undergoes assembly in response to enzymatic cleavage at Lys^4. [4, 10] This facilitates the formation of an inter-dimer β-strand, which enables the assembly of single dimeric units into HMW oligomers. The system, first described by Summers *et al.*, was successfully reproduced in this study, with a protein tecton forming upon proteolysis with 1:500 (w/v) trypsin:Nterm-Lsr2. However, using trypsin to initiate the assembly of Nterm-Lsr2 brought with it inherent problems. Unwanted proteolysis at additional lysine and arginine residues within the protein meant that the system was never optimal. High initial protein concentrations were needed to assemble large structures, and the identification of the oligomeric state was made difficult by the inclusion of hydrolysed products into the assemblies. This was seen both in solution (SEC) and on TEM, with the large "spaghetti-like" structures being masked by a significant amount of protein aggregate. The assembly system was significantly improved with the addition of a specific enzyme recognition site. Enteropeptidase digest was used to create a protein tecton that assembled into complex "spaghetti- like" structures at pH 7.4. Characterisation of the Ent-Nterm tecton showed that it was the same single unit that was formed during trypsin cleavage. Without the unwanted proteolysis, the system was able to perform at optimal levels.

6.1.2 Controlling tecton assembly

Through studying variations in environmental conditions it was possible to exert additional control over the assembly system. The sizes of the oligomers increased with increasing protein concentration. This was unsurprising for two reasons. First, the concentration used during the initial trials in this study was lower than that used by Summer *et al*, [4] resulting in smaller oligomers consistent with a concentration dependence on oligomer size. [10] The second reason is related to the mode of assembly. Each tecton has two oligomerisation sites located on either side of the 2-fold symmetry axis. The assembly system is "open" as a result and, with no capping mechanism, the oligomerisation can continue up to the point of protein saturation. [3] From the concentration variation experiments a model of assembly was

proposed whereby the LMW (tetramer) and MMW oligomers were in fast equilibrium, and the HMW oligomer only formed once the MMW oligomer population had reached a critical concentration. Additional equilibrium experiments corroborated this hypothesis.

Increasing the pH from 7.4 to 9.0 caused the oligomers to disassemble into a tetramer. This was attributed to the deprotonation of the N-terminal lysine α-amine (pK_a 8.95), which had been identified in the crystal structure as forming salt-bridges with negatively charged residues, and backbone carbonyl oxygen atoms on the neighbouring dimer (fig 6.1 B). It was hypothesised that the removal of these interactions resulted in a less stable interface. Protein concentrations higher than those used in this study would therefore be needed to drive oligomerisation into the HMW form. Further concentration analysis, and a crystal structure of the tecton at pH 9.0, could definitively confirm this hypothesis.

6.2 Controlled assembly of *Hs*Prx3-6his

*Hs*Prx3 is emerging as a useful building block for the formation of protein nanomaterials. [11, 12] This family of proteins has the ability to assemble into many different oligomeric states in response to environmental variations. [13-15]

6.2.1 *Hs*Prx3-6his is a stable protein tecton

For the typical 2-Cys Prx1 subfamily, the switch between dimer and toroid (decamer or dodecamer) usually occurs upon formation of the active site disulfide bond (Cys_p – Cys_r). [16-18] However, the addition of the histidine tag in this study resulted in a marked increase in stability of the dodecamer form. The active site disulfide bond was still able to form, as visualised by the SDS stable dimer, indicating that additional variations in the protein structure were the cause of the stability. When comparing the crystal structure of *Bt*Prx3 with the cryo-EM reconstitution of *Hs*Prx3 tubes at pH 4.0, it was noted that there were significant differences between geometries of the structures. [12] The root mean square deviation (RMSD) of residues at the A and B-type interface were measured. RMSD at the A-type interface was significantly greater than at the B-type interface (1.3 Å and 2.1 Å respectively). [12] The *Bt*Prx3 contains an N-terminal histidine tag and is also stable under oxidation

conditions. [19, 20] Additionally, the placement of residues at the A-type interface is important for toroid stabilisation. [18, 21, 22] Specific differences between the *Hs*Prx3-6his and *Hs*Prx3-WT at this interface could be the origin of the toroid stability. To confirm this, a crystal structure of the *Hs*Prx3-6his would need to be acquired under both reducing and non-reducing conditions. The A-type interface would be studied to show any variations that may occur in response to the active site disulfide bond formation. A further structure of the native protein (without the tag) would then be compared to the tagged structures to note any significant differences between the two.

Histidine residues at the A-type interface of *Ace*Prx1 are involved in the native decamer stability. [23] The imidazole side chains of His^{107} can interact via π-π stacking of the aromatic rings, along with H-bonding to backbone oxygen atoms on the adjacent monomer. This particular residue has been identified in another, similar, stable decameric form of Prx (*Hs*Prx4). [24] It is possible that the histidine residues within the current system (*Hs*Prx3-6his) are engaging in similar interactions. The crystal structure would be useful to corroborate this hypothesis, although the interactions may be dynamic in nature, as seen by the flexibility of the tag and linker in the bovine structure. [19] Additionally, the original histidine tag variation constructs, *Hs*Prx3-4his and *Hs*Prx3-2his, could be used to analyse this hypothesis. A reduction in toroid stability with decreasing histidine content would indicate that the toroid stabilisation was in response to direct interactions between the histidine residues and protein/additional tags within the ring.

6.2.2 Tubes and stacks formed within the physiological pH range

The formation of a stable toroid resulted in an ideal tecton from which larger assemblies could form. It has been noted in many studies that typical 2-Cys Prxs have the propensity to form stacks at the R-type interface. [12, 25-29] In fact, their geometry and symmetry are remarkably well suited for the formation of tubes. [30] There are a number of complementary residues that appear to enable this interaction, namely the backbone carbonyl groups on α2 with the protonated lysine and histidine residues on α6, as well as interactions between lysine and aspartic acid of β2 on adjacent subunits. [12, 31] It has been suggested that the HMW stabilisation may be due to changes in the environment and a switch in the enzyme function. [26, 27] However, some studies have noted that it may be possible for Prxs to interact at the

R-type interface in a dynamic manner at physiological pH, and in the peroxidase active form. [32]

With this is mind, the influence of the tag and linker sequence on the oligomeric state of *Hs*Prx3 was investigated. The most notable difference between this construct and the WT recombinant construct was that the pH sensitivity increased dramatically. The tubes, seen at pH 4.0 with *Hs*Prx3-WT, [11] began to form at much higher pH values with *Hs*Prx3-6his. A MW corresponding to two or three rings was seen at pH 7.6 on SEC and AUC, with TEM confirming that these were indeed stacks of Prx and not random aggregates. At pH 7.4, the rings had associated into tubes. These tubes were stable in solution with the oligomers eluting at V_0 on gel filtration, showing structures greater that the column MW limit (660 kDa). AUC further confirmed that the oligomers were stable in solution, and TEM analysis showed that the tubes were unusually uniform, when compared to other Prx tubes. [11, 33, 34] The length was further influenced by small changes in pH values. At pH 7.2 elongated tubes formed, with the most populated MW being around 5 x dodecamer in solution (AUC), increasing to 15 x dodecamer on TEM. The concentration was shown to have a notable effect on the MW of *Hs*Prx3-6his at pH 7.2, with the size increasing from 0.5 mg/mL to 5 mg/mL. This suggests a similar open helical symmetry to the tubes formed from *Hs*Prx3-WT at pH 4.2. [12] This could also be the reason why longer tubes were always seen on TEM, with the concentration increasing rapidly upon drying during TEM grid preparation.

Like many Prxs, the primary interactions at the R-type interface of are electrostatic in nature. [12, 26, 35] The tubes at pH 7.4 were subjected to a range of salt concentrations (NaCl) and analysed on SEC. As expected for electrostatic interactions the salt had a detrimental effect, with the HMW oligomers dissociating at 600 mM NaCl and above. The electrostatic interactions that had been noted in the native structure [12] were probed by substituting Lys^{23}, located on β2 at the R-type interface, with either His, Arg or Ala. While substitution for a neutral residue (*Hs*Prx3-K23A) led to a decrease in the proportion of HMW oligomers the association was not completely disrupted, suggesting the involvement of additional residues.

The points of potential charge within the non-native sequence are the histidine residues in the tag. Increasing the number of histidine residues (PRX3E and PRX3F, with 8 and 10 histidines respectively) led to an increase in the size of the assemblies, both in solution studies and on TEM. While the exact pK_a of these residues is unknown, it is possible to postulate a route

towards assembly. Histidine residues within a tag have varying pK_a values, depending on their proximity to the N-terminal amine. [36] It was hypothesised that the number of charged histidine residues within the tag increases with decreasing pH. The positively charged tag would be able to interact with the large area of negatively charged backbone carbonyl oxygen atoms on α2 of the adjacent ring. [26] This type of interaction has been seen at the A-type interface of other typical 2-Cys Prxs. [37] It has been suggested that Prxs rings have the potential to dynamically interact without the need for native structural changes in response to extreme environmental variations (for example high H_2O_2 concentration, [29] or acidic pH [11]). [32] As these interactions are normally short lived, and therefore not often noted, it is likely that the addition of the tag provides the stability needed for the tubes to form. To confirm this hypothesis, it would be beneficial to calculate the pK_a values of the histidine residues in the tag. With recent developments in cryo-EM allowing for the solution of structures to the near atomic level, [38] it may also be possible to elucidate the structure of the tubes formed from *Hs*Prx3-6his. [12] This could aid the identification of the exact points of contact at the R-type interface, which would likely confirm the mode of assembly.

6.2.3 Metal ion coordination

Metal ion coordination was also utilised in this study to initiate the onset of protein assembly. The imidazole side chains in the tag presented the ideal ligand to coordinate to divalent metal ions. [33, 39-41] It was found that Ni^{2+} produced the most stable oligomer, which is unsurprising considering its increased affinity for imidazole when compared to the other metals studied (Mg^{2+} and Co^{2+}). [39] The largest oligomer seen under all conditions was stacks of four dodecamers. This was less than had been seen in previous studies using a similar mode of assembly. [33] It was noted during the course of this study that the Prx used in the previous study, *Sm*Prx1, had additional metal chelating residues within the tag (the aspartic acid residues within the enteropeptidase cleavage site). One of the constructs designed during this project also had the same sequence of residues in the tag (PRX3D). Due to time constraints it was not possible to examine this protein under metal chelating conditions. This remains an area of potentially informative future work.

6.2.4 The length of the linker influences the oligomeric state

The influence of the length of the linker was also examined. This was due to the hypothesis that a disordered peptide within the toroid lumen may be instigating a chaperone response. Prxs have the ability to behave as holdase chaperones, preventing the aggregation of unfolded and partially unfolded proteins in response to cell stress conditions. [29, 35, 42] *Hs*Prx3-WT had been identified as a putative chaperone due to the large hydrophobic surface area within the toroid lumen. [12] Variations in the linker length did influence the quaternary structure of the protein, with stacks forming at pH 8.0 non-reducing conditions. The proteins analysed included the presence of an additional 6-histidine tag and linker sequence that was not within the original construct design. It is important to note that the increased number of histidines within the tag may have aided additional interactions between the linker sequence and the native protein (π-π stacking, H-bonding). That said, a significant number of Prx rings were visualised with electron density within the dodecamer on TEM. This was reminiscent of other Prx holdase chaperones. [42] Furthermore, the construct with the lowest peroxidase activity, PRX3B, also had the largest proportion of these chaperone-like species on TEM. It is possible that the drop in peroxidase activity was related to a switch in function, as there was no apparent trend in the activity and the degree of assembly. A crystal structure of these constructs would be useful in determining the exact reason behind the stacking. This may also provide insight into the loss in peroxidase activity. In addition, the structure could be compared to other known Prx chaperones to assess any similarities. A chaperone activity assay could also be completed to investigate whether there is a trend between the peroxidase and chaperone activity of these constructs.

6.3 Functionalisation opportunities for the novel protein building-blocks

The aim of this study was to develop novel assembly systems to be used in the formation of functional biological nanomaterials. As a result of the work on Nterm-Lsr2, it was possible to design proteins that assembled upon a specific cleavage event. The enteropeptidase recognition site is found in many recombinant expression systems, [43, 44] and is already encoded into the protein sequence from the expression plasmid. This makes it an ideal target for use in other protein assembly systems, providing a means for oligomerisation on demand.

Ent-Nterm assembles into long "spaghetti-like" structures. The amino acid content of these structures means there is a wealth of surface functionalisation sites. [45] Using the methods developed by Sasso *et al*. (Chapter 1 at 1.4.3) it may be possible to add functional entities to these structures. These functional assemblies could be used in a range of different fields, from biosensors to high surface area catalysts. [45], [46]

The water soluble tubes formed from the human protein *Hs*Prx3-6his could be developed for the use in biomedical applications. Protein nanotubes are gaining interest within this field due their open-ended structure (which is ideal for drug delivery and storage), the longevity in which they can exist in the body, their biocompatibility, and the large surface area available for functionalisation on both the interior and exterior of the tube. [47] The stability of *Hs*Prx3-6his 1D tubes at pH values around physiological pH values lends further credence to their potential use within this field.

Proteins have been shown to rapidly adsorb on gold surfaces due to interactions between Au and cysteine, aspartic acid, and glutamic acid residues found on the surface of many proteins. [48] While the cysteine residues in *Hs*Prx3-6his are buried within the protein interior, there are still a number of available residues that can interact with gold. [49] The system of *Hs*Prx3-6his HMW oligomerisation, developed during this study, has been used to investigate the interaction of the protein with Au (111) gold terraces. (Laura Domigan, unpublished work). Scanning tunnelling electron microscopy (STEM) was used to image the arrangement of the protein, which bound readily to the surface (fig 6.1 A).

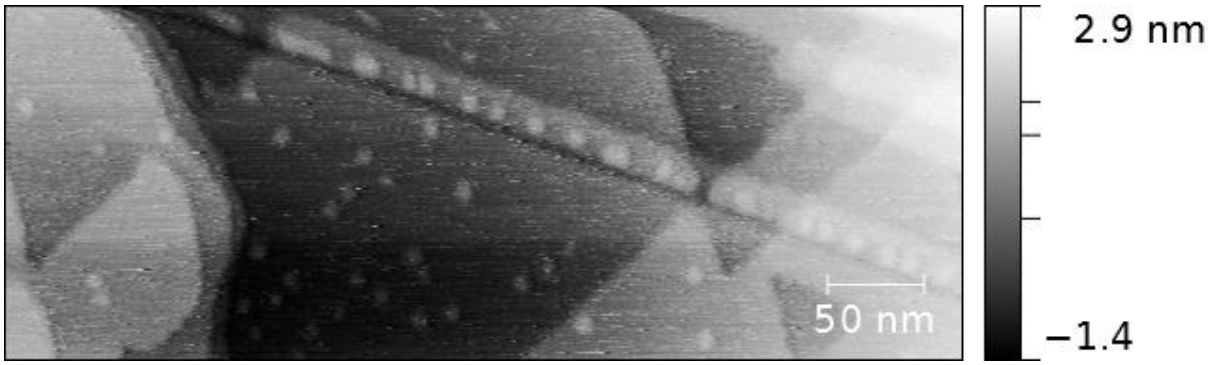


Figure 6.1: STEM of *Hs*Prx3-6his bound to a Au(111) surface at pH 7.2. The protein, at ~ 2 nm in height along the D6 axis, is aligned as an array of rings upon the gold surface.

The influence of pH on the assembly was further probed using quartz crystal microbalance with dissipation (QCM-D). This technique measures the resonance frequency and dissipation in response to adsorption onto a sensor surface. [50, 51] The mass of adsorbed particles is given by the change in the frequency, and changes in dissipation provide information about the viscoelasticity of the absorbed layer. The mass calculated from ΔF showed the adsorption of a monolayer of *Hs*Prx3-6his at pH 8.0 (fig. 6.2 i), and a greater adsorption mass at pH 7.2 due to the oligomers formed at this pH (fig. 6.2 ii). The dissipation was also greater for the protein at pH 7.2 (fig. 6.2 i and ii). It is believed this is due the heterogeneous nature of these oligomers, leading to a less ridged and structured surface layer.

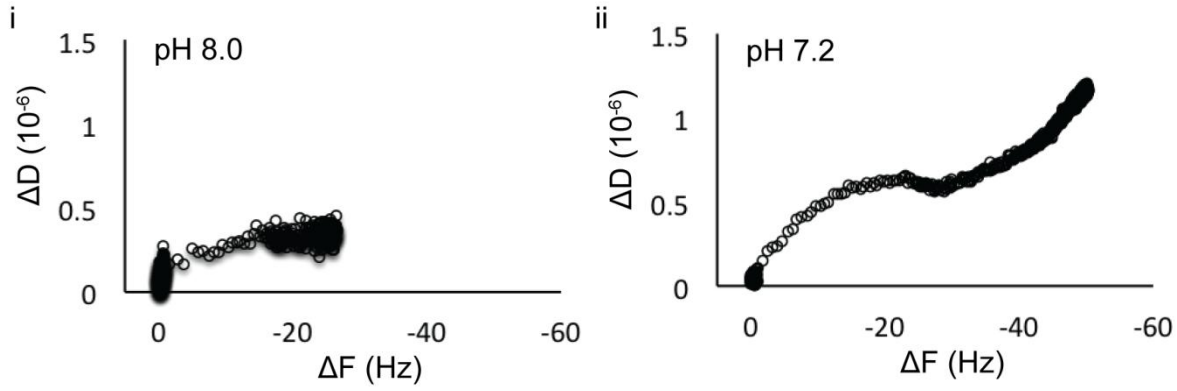


Figure 6.2: Change of ΔD dissipation plotted as a function of change in frequency ΔF at (i) pH 8.0 and (ii) pH 7.2. The change in frequency is less at pH 8.0, showing a lower mass of adsorbed protein. At pH 7.2 the adsorption appears to occur in two stages. The first stage is the protein binding to the gold, and the second stage is the oligomerisation of the protein in response to the low pH.

The results from this study provide a proof of principle that the sizes and properties of *Hs*Prx3-6his layers on a gold surface can be manipulated by the pH sensitivity. The control over the adsorption could aid the formation of ordered protein arrays on a gold surface. The metal bound tags within the lumen could be further utilised to bind to additional nanoparticles to form complex nanoscale systems with a variety of different functions.

6.4 References

1. Yeates T: Nanobiotechnology: protein arrays made to order. *Nature Nanotechnology* 2011, 6:541 - 542.
2. Yang L, Zhang LJ, Webster TJ: Nanobiomaterials: State of the art and future trends. *Advanced Engineering Materials* 2011, 13(6):B197-B217.
3. Lai YT, King NP, Yeates TO: Principles for designing ordered protein assemblies. *Trends in Cell Biology* 2012, 22(12):653-661.
4. Summers EL, Meindl K, Uson I, Mitra AK, Radjainia M, Colangeli R, Alland D, Arcus VL: The structure of the oligomerization domain of Lsr2 from Mycobacterium tuberculosis reveals a mechanism for chromosome organization and protection. *PloS ONE* 2012, 7(6).
5. Gordon BRG, Li YF, Wang LR, Sintsova A, van Bakel H, Tian SH, Navarre WW, Xia B, Liu J: Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America* 2010, 107(11):5154-5159.
6. Chen JM, Ren H, Shaw JE, Wang YJ, Li M, Leung AS, Tran V, Berbenetz NM, Kocincova D, Yip CM *et al*: Lsr2 of Mycobacterium tuberculosis is a DNA-bridging protein. *Nucleic Acids Research* 2008, 36(7):2123-2135.
7. Ali SS, Xia B, Liu J, Navarre WW: Silencing of foreign DNA in bacteria. *Current Opinion in Microbiology* 2012, 15(2):175-181.
8. Qu Y, Lim CJ, Whang YR, Liu J, Yan J: Mechanism of DNA organization by *Mycobacterium tuberculosis* protein Lsr2. *Nucleic Acids Research* 2013, 41(10):5263-5272.
9. Colangeli R, Haq A, Arcus VL, Summers E, Magliozzo RS, McBride A, Mitra AK, Radjainia M, Khajo A, Jacobs WR *et al*: The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proceedings of the National Academy of Sciences of the United States of America* 2009, 106(11):4414-4418.
10. Ashmead HM, Negron L, Webster K, Arcus V, Gerrard JA: Proteins as supramolecular building blocks: Nterm-Lsr2 as a new protein tecton. *Biopolymers* 2015, 103(5):260-270.
11. Phillips AJ, Littlejohn J, Yewdall NA, Zhu T, Valery C, Pearce FG, Mitra AK, Radjainia M, Gerrard JA: Peroxiredoxin is a versatile self-assembling tecton for protein nanotechnology. *Biomacromolecules* 2014, 15(5):1871-1881.
12. Radjainia M, Venugopal H, Desfosses A, Phillips AJ, Yewdall NA, Hampton MB, Gerrard JA, Mitra AK: Cryo-Electron Microscopy structure of human peroxiredoxin-3 filament reveals the assembly of a putative chaperone. *Structure* 2015, 23(5):912-920.
13. Perkins A, Poole LB, Karplus PA: Tuning of peroxiredoxin catalysis for various physiological. *Biochemistry* 2014, 53(49):7693-7705.
14. Wood ZA, Poole LB, Karplus PA: Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science* 2003, 300(5619):650-653.
15. Karplus PA: A primer on peroxiredoxin biochemistry. *Free Radical Biology and Medicine* 2015, 80:183-190.
16. Leslie BP, and Kimberly JN: Distribution and features of the six classes of Peroxiredoxins. *Molecular Cell* 2016, 39(1):53-59.
17. Wood ZA, Schroder E, Harris JR, Poole LB: Structure, mechanism and regulation of peroxiredoxins. *Trends in Biochemical Sciences* 2003, 28(1):32-40.

18. Matsumura T, Okamoto K, Iwahara S-I, Hori H, Takahashi Y, Nishino T, Abe Y: Dimer-oligomer interconversion of wild-type and mutant rat 2-Cys peroxiredoxin. *Journal of Biological Chemistry* 2008, 283(1):284-293.
19. Cao Z, Bhella D, Lindsay JG: Reconstitution of the mitochondrial PrxIII antioxidant defence pathway: General properties and factors affecting PrxII activity and oligomeric state. *Journal of Molecular Biology* 2007, 372(4):1022-1033.
20. Cao ZB, Roszak AW, Gourlay LJ, Lindsay JG, Isaacs NW: Bovine mitochondrial Peroxiredoxin III forms a two-ring catenane. *Structure* 2005, 13(11):1661-1664.
21. Wood ZA, Poole LB, Hantgan RR, Karplus PA: Dimers to doughnuts: Redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry* 2002, 41(17):5493-5504.
22. Morais MAB, Giuseppe PO, Souza TACB, Alegria TGP, Oliveira MA, Netto LES, Murakami MT: How pH modulates the dimer-decamer interconversion of 2-Cys Peroxiredoxins from the Prx1 subfamily. *Journal of Biological Chemistry* 2015, 290(13):8582-8590.
23. Nguyen JB, Pool CD, Wong CYB, Treger RS, Williams DL, Cappello M, Lea WA, Simeonov A, Vermeire JJ, Modis Y: Peroxiredoxin-1 from the human hookworm Ancylostoma ceylanicum forms a stable oxidized decamer and is covalently inhibited by conoidin A. *Chemistry & Biology* 2013, 20(8):991-1001.
24. Cao ZB, Tavender TJ, Roszak AW, Cogdell RJ, Bulleid NJ: Crystal structure of reduced and of oxidized peroxiredoxin IV enzyme reveals a stable oxidized decamer and a non-disulfide-bonded intermediate in the catalytic cycle. *Journal of Biological Chemistry* 2011, 286(49):42257-42266.
25. Perkins A, Nelson KJ, Parsonage D, Poole LB, Karplus PA: Peroxiredoxins: guardians against oxidative stress and modulators of peroxide signaling. *Trends in Biochemical Sciences* 2015, 40(8):435-445.
26. Saccoccia F, Di Micco P, Boumis G, Brunori M, Koutris I, Miele AE, Morea V, Sriratana P, Williams DL, Bellelli A *et al*: Moonlighting by different stressors: Crystal structure of the chaperone species of a 2-Cys Peroxiredoxin. *Structure* 2012, 20(3):429-439.
27. Kumsta C, Jakob U: Redox-regulated chaperones. *Biochemistry* 2009, 48(22):4666-4676.
28. Hall A, Karplus PA, Poole LB: Typical 2-Cys peroxiredoxins - structures, mechanisms and functions. *Febs Journal* 2009, 276(9):2469-2477.
29. Jang HH, Lee KO, Chi YH, Jung BG, Park SK, Park JH, Lee JR, Lee SS, Moon JC, Yun JW *et al*: Two enzymes in one: Two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell* 2004, 117(5):625-635.
30. Angelucci F, Bellelli A, Ardini M, Ippoliti R, Saccoccia F, Morea V: One ring (or two) to hold them all - on the structure and function of protein nanotubes. *Febs Journal* 2015, 282(15):2827-2845.
31. Saccoccia F, Angelucci F, Boumis G, Desiato G, Miele AE, Bellelli A: Selenocysteine robustness versus cysteine versatility: a hypothesis on the evolution of the moonlighting behaviour of peroxiredoxins. *Biochemical Society Transactions* 2014, 42(6):1768-1772.
32. Noichri Y, Palais G, Ruby V, D'Autreaux B, Delaunay-Moisan A, Nystrom T, Molin M, Toledano MB: In vivo parameters influencing 2-Cys Prx oligomerization: The role of enzyme sulfinylation. *Redox Biology* 2015, 6:326-333.
33. Ardini M, Giansanti F, Di Leandro L, Pitari G, Cimini A, Ottaviano L, Donarelli M, Santucci S, Angelucci F, Ippoliti R: Metal-induced self-assembly of Peroxiredoxin as

a tool for sorting ultrasmall gold nanoparticles into one-dimensional clusters. *Nanoscale* 2014, 6(14):8052-8061.

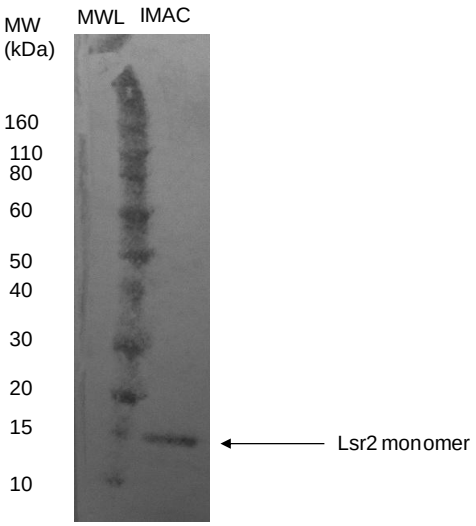
34. Gourlay LJ, Bhella D, Kelly SM, Price NC, Lindsay JG: Structure-function analysis of recombinant substrate protein 22 kDa (SP-22) - A mitochondrial 2-Cys peroxiredoxin organized as a decameric toroid. *Journal of Biological Chemistry* 2003, 278(35):32631-32637.
35. Angelucci F, Saccoccia F, Ardini M, Boumis G, Brunori M, Di Leandro L, Ippolitl R, Miele AE, Natoli G, Scotti S *et al*: Switching between the alternative structures and functions of a 2-Cys peroxiredoxin, by site-directed mutagenesis. *Journal of Molecular Biology* 2013, 425(22):4556-4568.
36. Valenti LE, De Pauli CP, Giacomelli CE: The binding of Ni(II) ions to hexahistidine as a model system of the interaction between nickel and His-tagged proteins. *Journal of Inorganic Biochemistry* 2006, 100(2):192-200.
37. Morais MAB, Giuseppe PO, Souza T, Alegria TGP, Oliveira MA, Netto LES, Murakami MT: How pH modulates the dimer-decamer interconversion of 2-Cys peroxiredoxins from the Prx1 subfamily. *Journal of Biological Chemistry* 2015, 290(13):8582-8590.
38. Yan Z, Bai XC, Yan CY, Wu JP, Li ZQ, Xie T, Peng W, Yin CC, Li XM, Scheres SHW *et al*: Structure of the rabbit ryanodine receptor RyR1 at near-atomic resolution. *Nature* 2015, 517(7532):50-+.
39. Bornhorst JA, Falke JJ: Purification of proteins using polyhistidine affinity tags. *Methods in Enzymology* 2000, 326:245-254.
40. Brodin JD, Carr JR, Sontz PA, Tezcan FA: Exceptionally stable, redox-active supramolecular protein assemblies with emergent properties. *Proceedings of the National Academy of Sciences of the United States of America* 2014, 111(8):2897-2902.
41. Bai YS, Luo Q, Zhang W, Miao L, Xu JY, Li HB, Liu JQ: Highly ordered protein nanorings designed by accurate control of gutathione S-transferase self-assembly. *Journal of the American Chemical Society* 2013, 135(30):10966-10969.
42. Teixeira F, Castro H, Cruz T, Tse E, Koldewey P, Southworth DR, Tomas AM, Jakob U: Mitochondrial peroxiredoxin functions as crucial chaperone reservoir in *Leishmania infantum*. *Proceedings of the National Academy of Sciences of the United States of America* 2015, 112(7):E616-E624.
43. Mikhailova AG, Rumsh LD: Enteropeptidase - Structure, function, and application in biotechnology. *Applied Biochemistry and Biotechnology* 2000, 88(1-3):159-174.
44. Arnau J, Lauritzen C, Petersen GE, Pedersen J: Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification* 2006, 48(1):1-13.
45. Sasso L, Suei S, Domigan L, Healy J, Nock V, Williams MAK, Gerrard JA: Versatile multi-functionalization of protein nanofibrils for biosensor applications. *Nanoscale* 2014, 6(3):1629-1634.
46. Komatsu Y, Fukuda M, Yamada H, Kawamoto S, Miyakawa T, Morikawa R, Takasu M, Yokojima S, Akanuma S, Yamagishi A: Constructing protein nanofiber and estimation of the electronic state around metal ions. *International Journal of Quantum Chemistry* 2012, 112(24):3750-3755.
47. Komatsu T: Protein-based nanotubes for biomedical applications. *Nanoscale* 2012, 4(6):1910-1918.
48. Welinder AC, Zhang J, Steensgaard DB, Ulstrup J: Adsorption of human insulin on single-crystal gold surfaces investigated by in situ scanning tunnelling microscopy

and electrochemistry. *Physical Chemistry Chemical Physics* 2010, 12(34):9999-10011.

49. Cao Z, McGow DP, Shepherd C, Lindsay JG: Improved catenated structures of bovine peroxiredoxin III F190L reveal details of ring-ring interactions and a novel conformational state. *PLOS One* 2015, 10(4).
50. Marx KA: Quartz crystal microbalance: A useful tool for studying thin polymer films and complex biomolecular systems at the solution-surface interface. *Biomacromolecules* 2003, 4(5):1099-1120.
51. Keller CA, Kasemo B: Surface specific kinetics of lipid vesicle adsorption measured with a quartz crystal microbalance. *Biophysical Journal* 1998, 75(3):1397-1402.

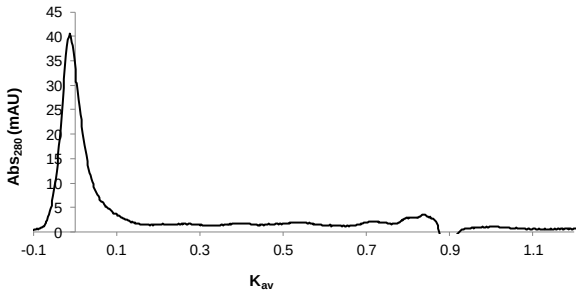
Appendix

Appendix: Chapter Three

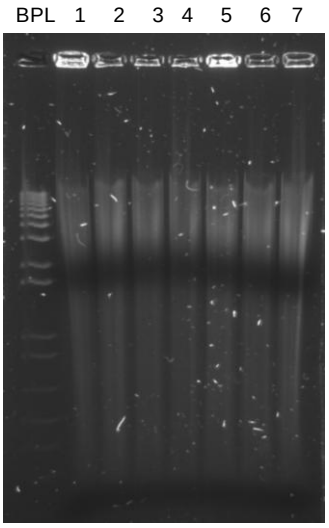


Appendix 3.1: Western blot of Lsr2 fraction for IMAC. Only visible band is the Lsr2 monomer indicating that other band seen on SDS-PAGE is an impurity that does not contain a histidine tag.

A: SEC of Lsr2, 1 M NaCl

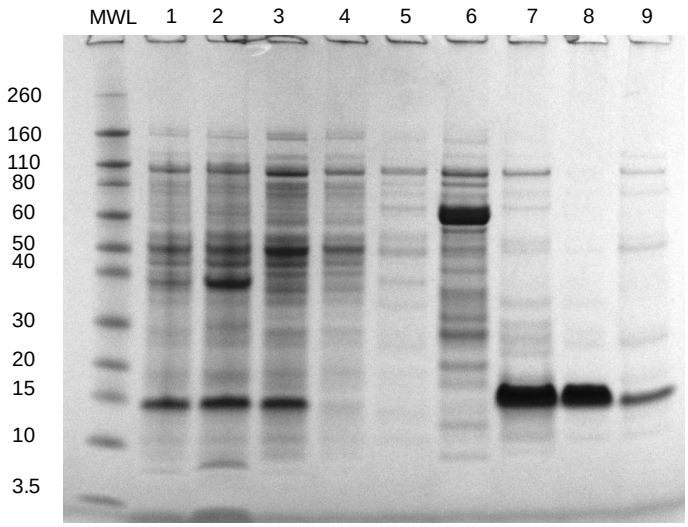


B: Agarose gel of DNAse digest

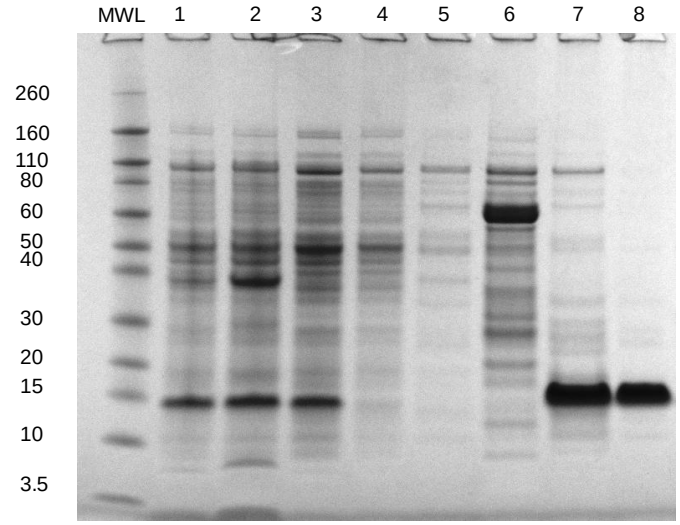


Appendix 3.2: Analysis of different conditions to try and remove DNA from full length Lsr2. BPL stands for base pair ladder. (A) SEC of Lsr2 in high salt buffer (50 mM NaP_i, 1 M NaCl, pH 7.4). Peak is still at the void showing that Lsr2 is still bound tightly to DNA. (B) agarose gel of DNAse digest.1, 2, 3, 4, 5, 6 and 7 - samples left at room temperature for 5, 10, 15, 30, 60, 120 and 240 minutes respectively. There is no visible digest of the genomic DNA implying that Lsr2 is protecting the DNA.

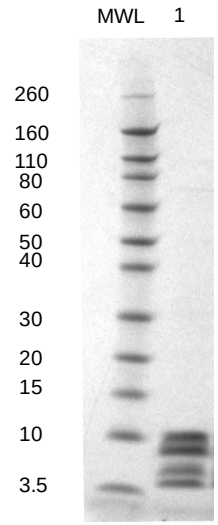
A:Lsr2-R97A



B: LSr2-R97AR99A

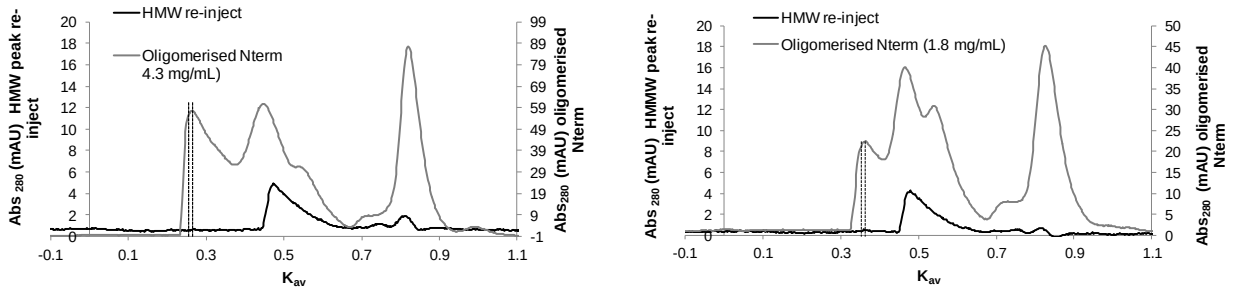


Appendix 3.3: SDS-PAGE analysis of DNA binding mutants. (A) Lsr2- R97A mutant agarose gel, 1 – crude lysate, 2 – insoluble crude, 3 – soluble crude, 4 – IMAC flow through, 5 – IMAC peak 1, 6 – IMAC peak 2, 7 – IMAC peak 3 (all IMAC peaks contain Lsr2-R97A). (B) Lsr2-R97AR99A mutant agarose gel, 1 – crude lysate, 2 – insoluble crude, 3 – soluble crude, 4 – IMAC flow through, 5 – IMAC peak 1, 6 – IMAC peak 2 (all IMAC peaks contain Lsr2-R97AR99A).



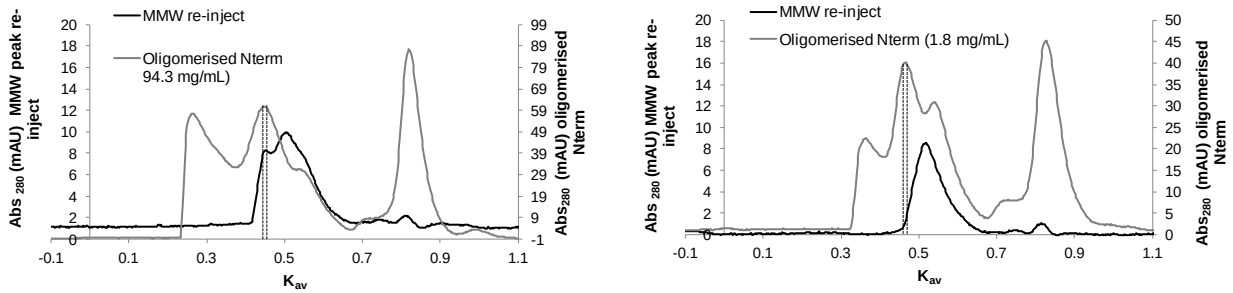
Appendix 3.4: SDS showing the fractionation of Nterm-Lsr2 due to unwanted trypsin cleavage. 1 - the protein fragment peak.

A: 4.3 mg/mL **B: 1.8 mg/mL**



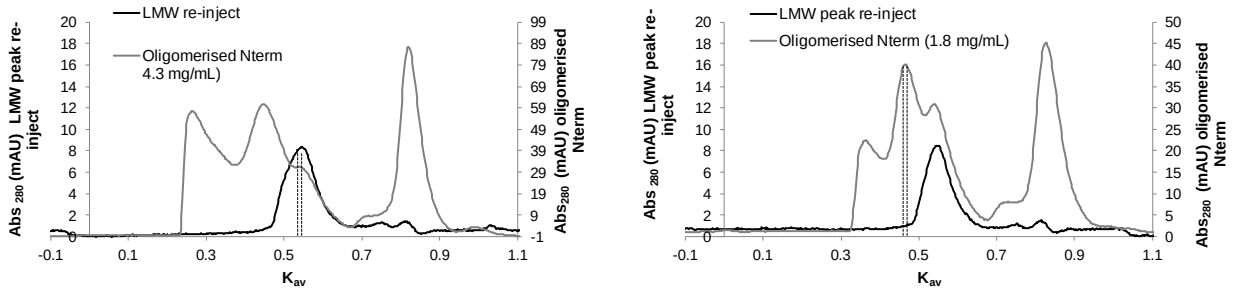
Appendix 3.5: Equilibrium of HMW assemblies of Nterm-Lsr2: (A) Equilibrium of 30-mer from 4.3 mg/mL oligomerisation of Nterm-Lsr2; (B) Equilibrium of tetradecamer from 1.8 mg/mL oligomerisation of Nterm-Lsr2.

A: 4.3 mg/mL **B: 1.8 mg/mL**

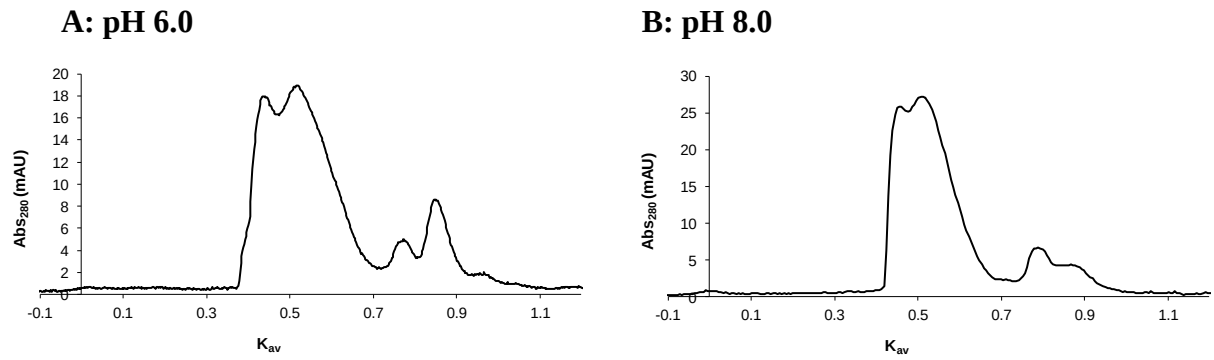


Appendix 3.6: Equilibrium of MMW assemblies of Nterm-Lsr2: (A) Equilibrium of octamer from 4.3 mg/mL oligomerisation of Nterm-Lsr2; (B) Equilibrium of heptamer from 1.8 mg/mL oligomerisation of Nterm-Lsr2.

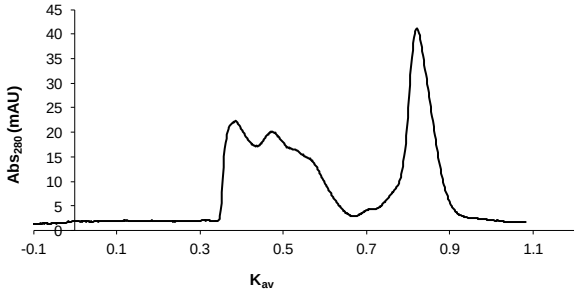
A: 4.3 mg/mL **B: 1.8 mg/mL**



Appendix 3.7: Equilibrium of LMW assemblies of Nterm-Lsr2. (A) Equilibrium of tetramer from 4.3 mg/mL oligomerisation of Nterm-Lsr2; (B) Equilibrium of tetramer from 1.8 mg/mL oligomerisation of Nterm-Lsr2.



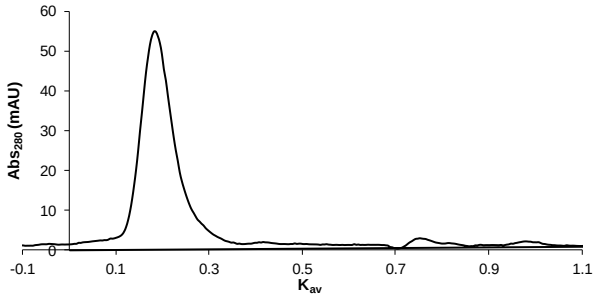
Appendix 3.8: Oligomerisation at varying pH. (A) pH 6.0 and (B) pH 8.0 There is a small change in the extent of oligomerisation with the MMW oligomer being the largest formed.



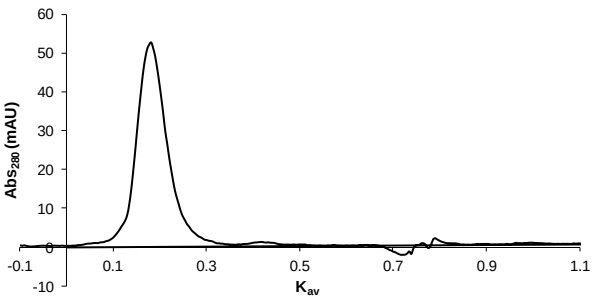
Appendix 3.9: Oligomerisation of Nterm-Lsr2 in tris buffer at pH 7.4. There is no significant difference in the oligomerisation within different buffer systems at pH 7.4.

Appendix: Chapter Four

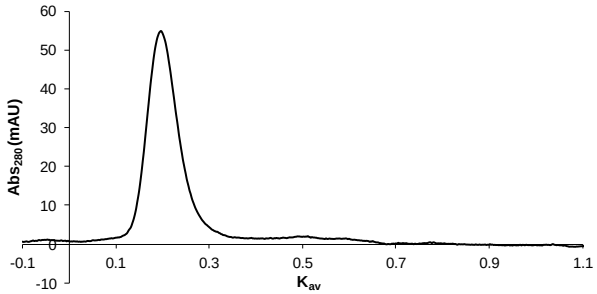
A: 10 mM EDTA O/N



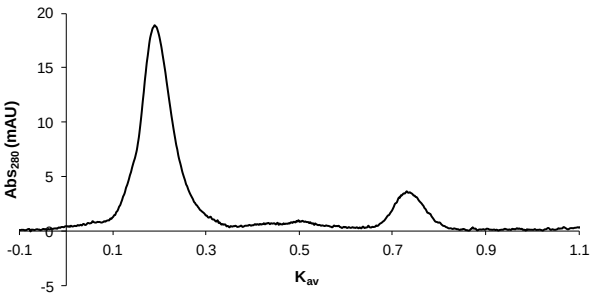
B: 50 mM EDTA O/N



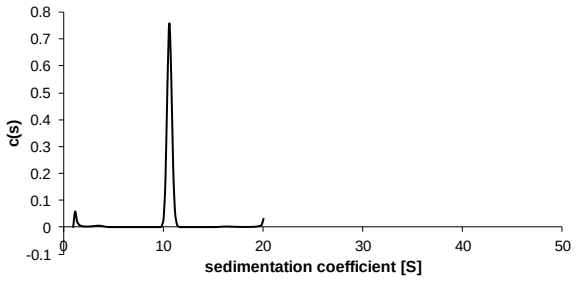
C: 10 mM EDTA + 20 mM imidazole



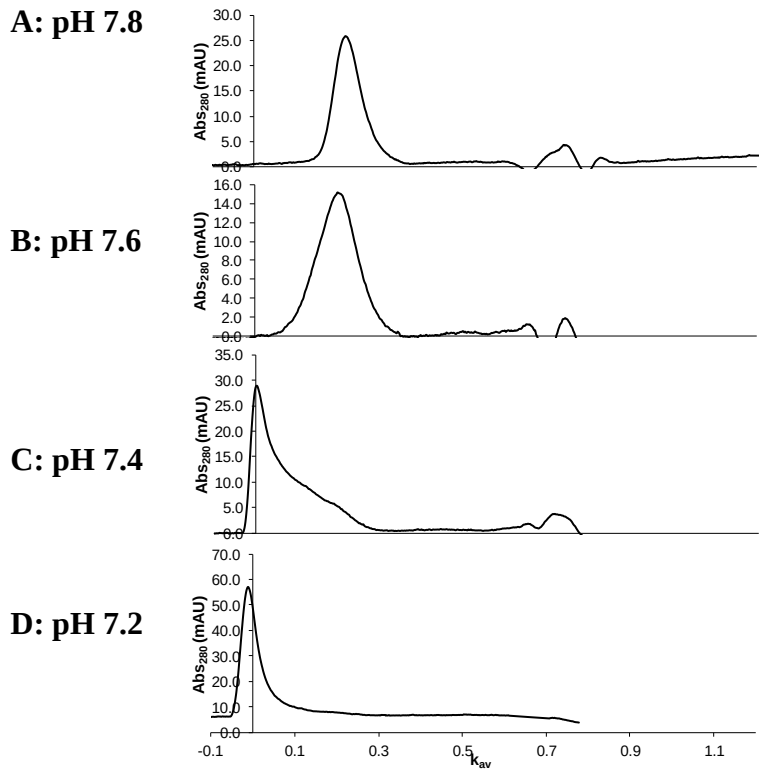
D: low protein concentration



Appendix 4.1: *Hs*Prx3-6his in non-reducing conditions. A: increased dissociation time from one hour to O/N in EDTA, b: Increased EDTA concentration, c: additional chelating agent EDTA and imidazole, d: low protein concentration (0.4 mg/mL). Non of the different conditions led to the dissociation of the decamer into dimers.

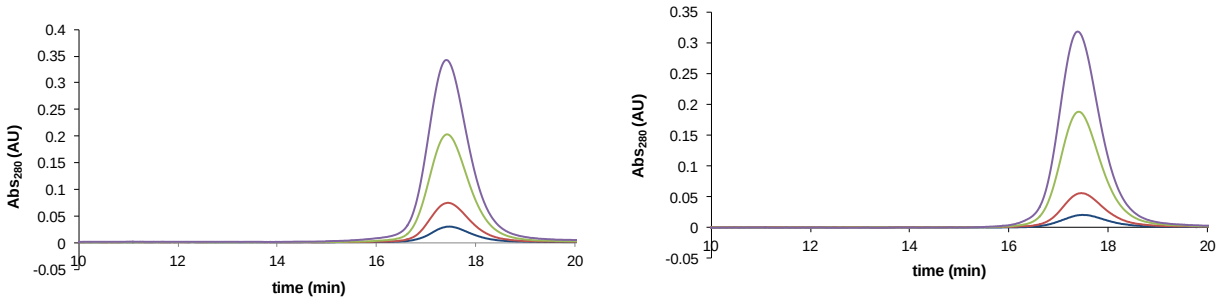


Appendix 4.2: AUC of *Hs*Prx3-6his in reducing conditions. There is a single species at the molecular weight of a dodecamer.

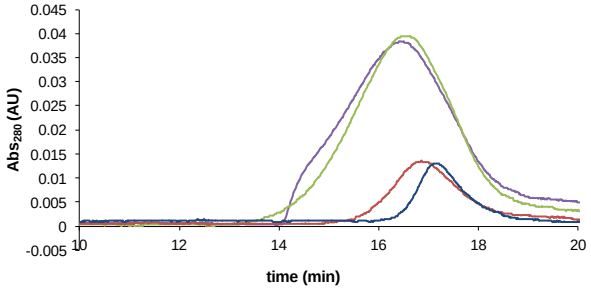


Appendix 4.3: SEC of oligomerisation of *Hs*Prx3-6his with decreasing pH in reducing conditions. (A) pH 7.8, (B) pH 7.6,(C) pH 7.4 and (D) pH 7.2. The switch between single rings and HMW oligomers occurs between pH 7.6 and 7.4. The oligomerisation is slightly less at pH 7.4 than in non-reduced conditions.

A: pH 7.8 **B: pH 7.6**

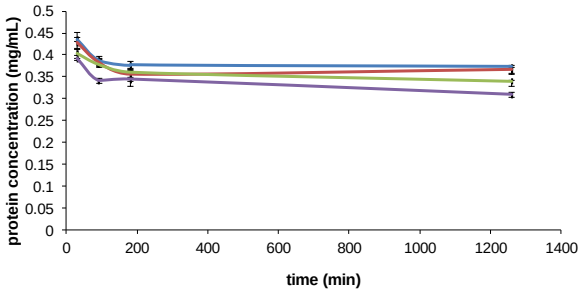


Appendix 4.4: HPLC trace of the concentration dependence on the size of oligomers at pH (A) 7.8 and (B) 7.6. blue: 0.5 mg/mL; red: 1 mg/mL, green: 3 mg/mL and purple : 5 mg/mL. The oligomer mass is the equivalent of a double stack at pH 7.6 but there is no significant concentration dependence on the oligomer size.



Appendix 4.5: HPLC trace of the concentration dependence on the size of oligomers at pH 7.2, blue: 0.5 mg/mL; red: 1 mg/mL, green: 3 mg/mL and purple : 5 mg/mL. Increasing the protein concentration leads to a significant increase in the oligomer mass.

A: NaCl



B: $(NH4)_2SO4$

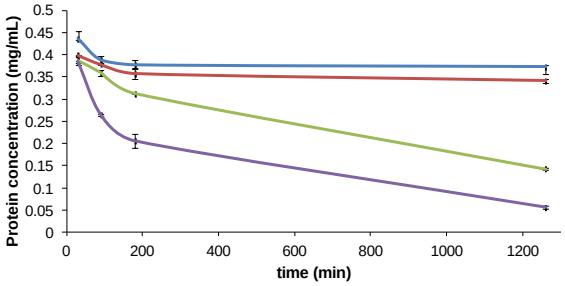
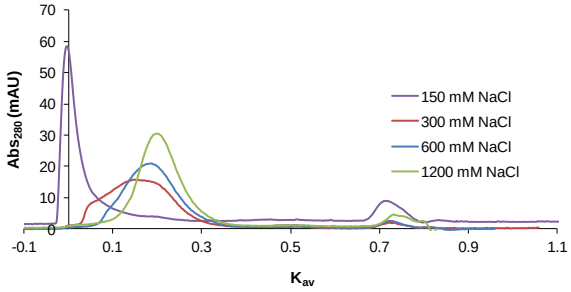
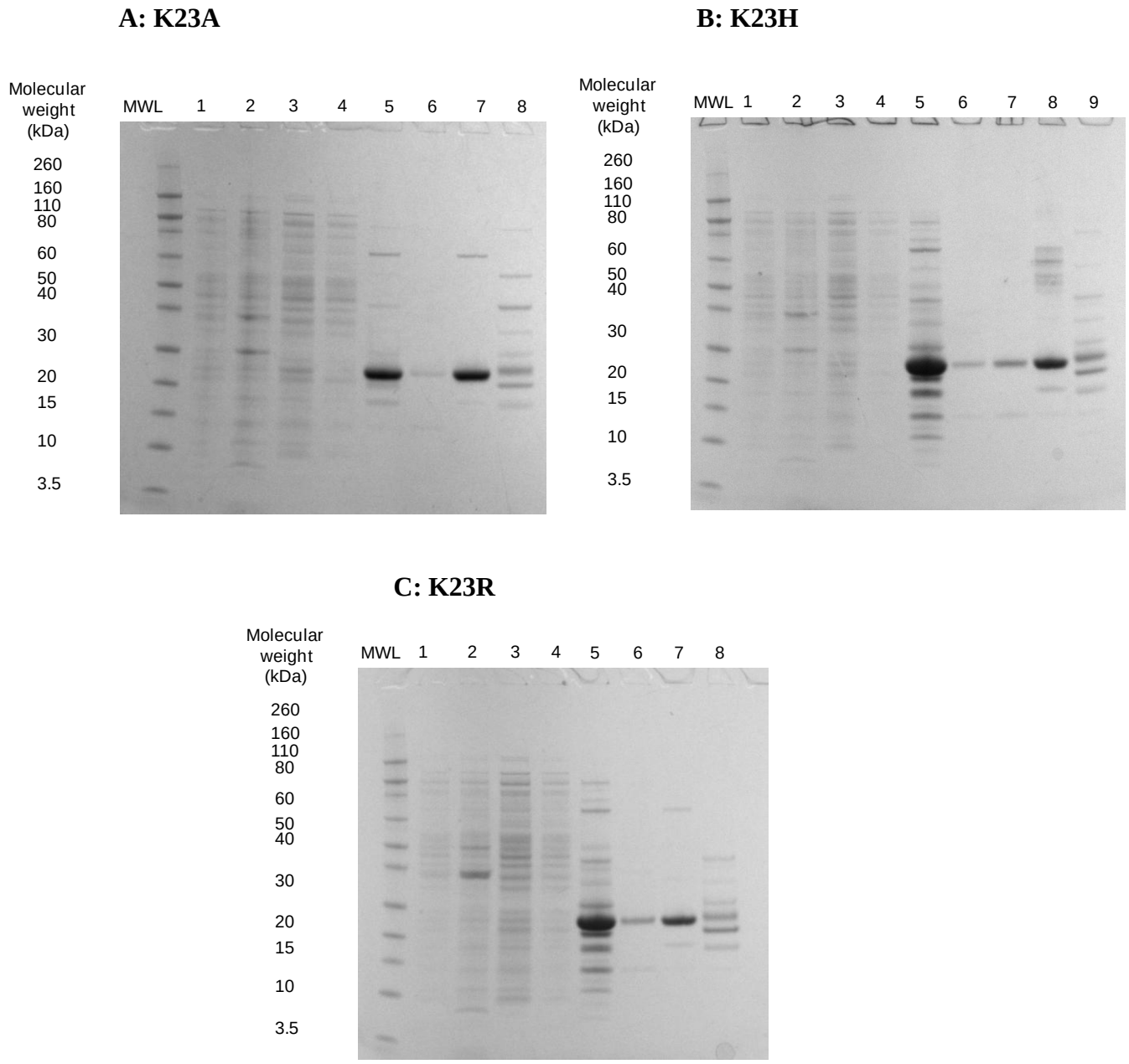


Figure 4.6: Influence of salt concentration on the concentration of *Hs*Prx3-6his over time. (A) NaCl and (B) $(NH_4)_2SO_4$. Blue: no salt; red: 100 mM; green: 200 mM and purple: 300 mM. While there is very little variation in the protein concentration with NaCl there is a significant decrease in concentration in the presence of > 100 mM $(NH_4)_2SO_4$. For this reason $(NH_4)_2SO_4$ was not used to try and disrupt tube formation so as to avoid false positives.



Appendix 4.7: Increasing salt concentration disrupts HMW oligomerisation. The oligomerisation is reduced at 300 mM NaCl and completely disrupted at 1.2 M NaCl.

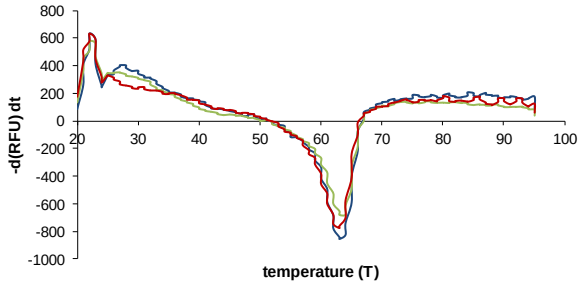


Appendix 4.7: SDS-PAGE analysis of the purification of the Lys[23] mutants. (A) *Hs*Prx3-K23A. 1: crude lysate; 2: insoluble crude 1; 3: soluble crude; 4: IMAC flow through; 5: IMAC imidazole eluted fraction; 6: SEC peak 1; 7: SEC peak 2; 8: SEC peak 3. (B) *Hs*Prx3-K23H. 1: crude lysate; 2: insoluble crude 1; 3: soluble crude; 4: IMAC flow through; 5: IMAC imidazole eluted fraction; 6: SEC peak 1; 7: SEC peak 2; 8: SEC peak 3, 9: SEC peak 4. (C) *Hs*Prx3-K23R. 1: crude lysate; 2: insoluble crude 1; 3: soluble crude; 4: IMAC flow through; 5: IMAC imidazole eluted fraction; 6: SEC peak 1; 7: SEC peak 2; 8: SEC peak 3.

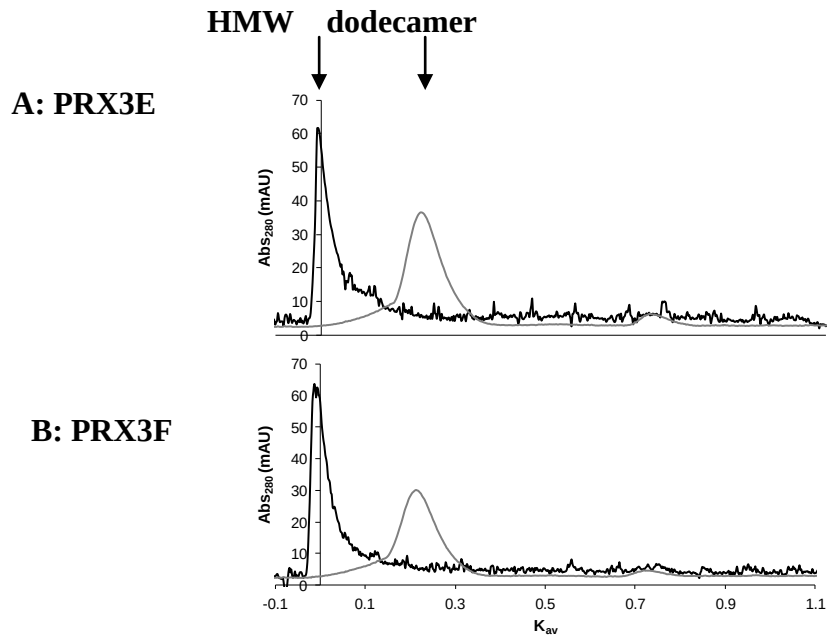
Appendix: Chapter Five

Protein	LCMS (m/z)	Calculated mass (kDa)
PRX3A	25309	26073
PRX3B	25950	26065
PRX3C	24530	24623
PRX3D	25528	25644
PRX3E	26805	26904
PRX3F	27801	27174

Appendix 5.1: Molecular weight of the tag and linker constructs as measured by LCMS

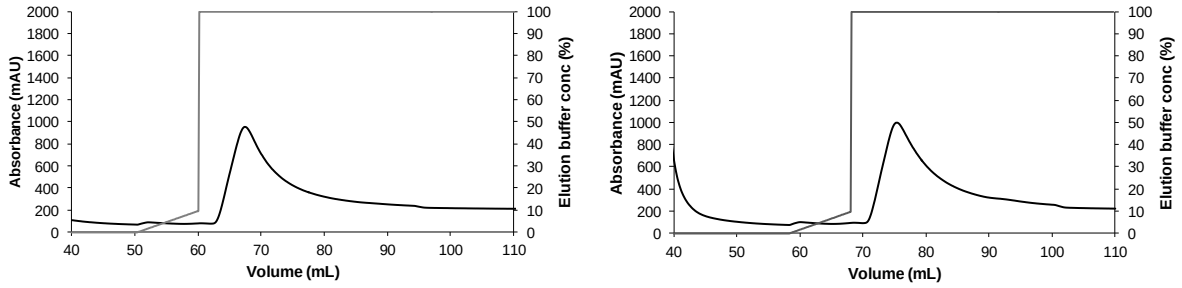


Appendix 5.2: Differential scanning fluorimetry of the histidine tag varying and *Hs*Prx3-6his: red: 6-histidine; blue: 10-histidine and green: 8-histidine. The relative fluorescent unite (RFU) increases with increasing unfolding as the dye (sypro-orange) binds to the hydrophobic regions of the protein. By plotting -d(RFU) dt against temperature it is possible to calculate the melting temperature of the protein. The melting temperature for all 3 proteins was 63 ^{o}C showing that there was no decrease or increase in the protein stability with the varying tag and linker length.

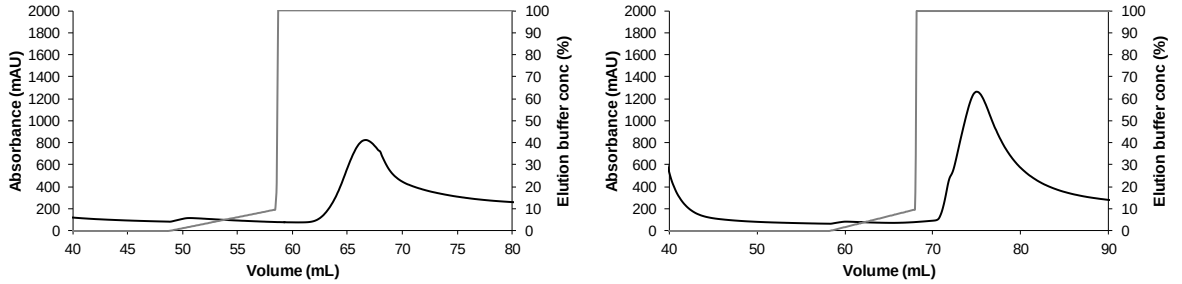


Appendix 5.3: SEC traces of 8-histidine (PRX3E) and 10-histidine tagged (PRX3F) *Hs*Prx3 at pH 7.4. The longer histidine tagged constructs, (A) PRX3E and (B) PRX3F, where dialysed into pH 7.4 buffer and run on SEC. This was to check if they had the same response to pH changes as *Hs*Prx3-6his. Both constructs assemble into HMW oligomers at pH 7.4.

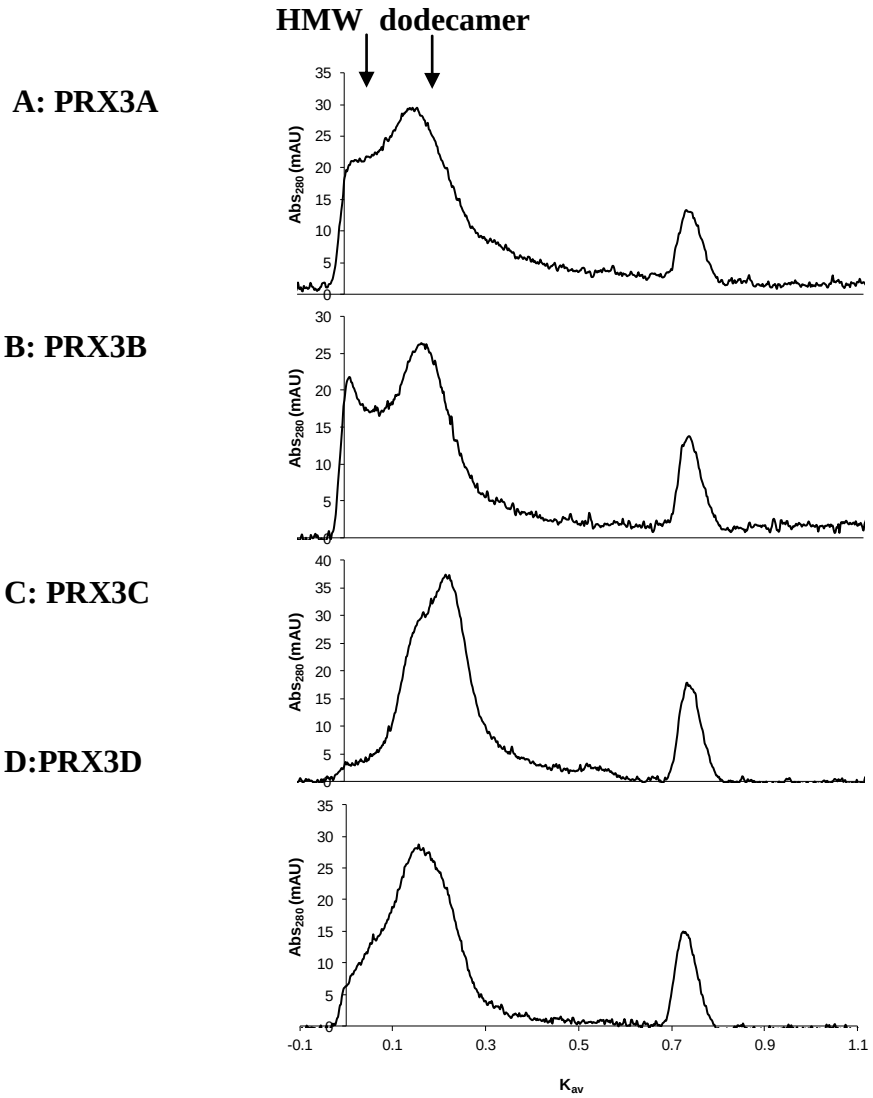
A: PRX3A IMAC **B: PRX3B IMAC**



C: PRX3C IMAC **D: PRX3D IMAC**

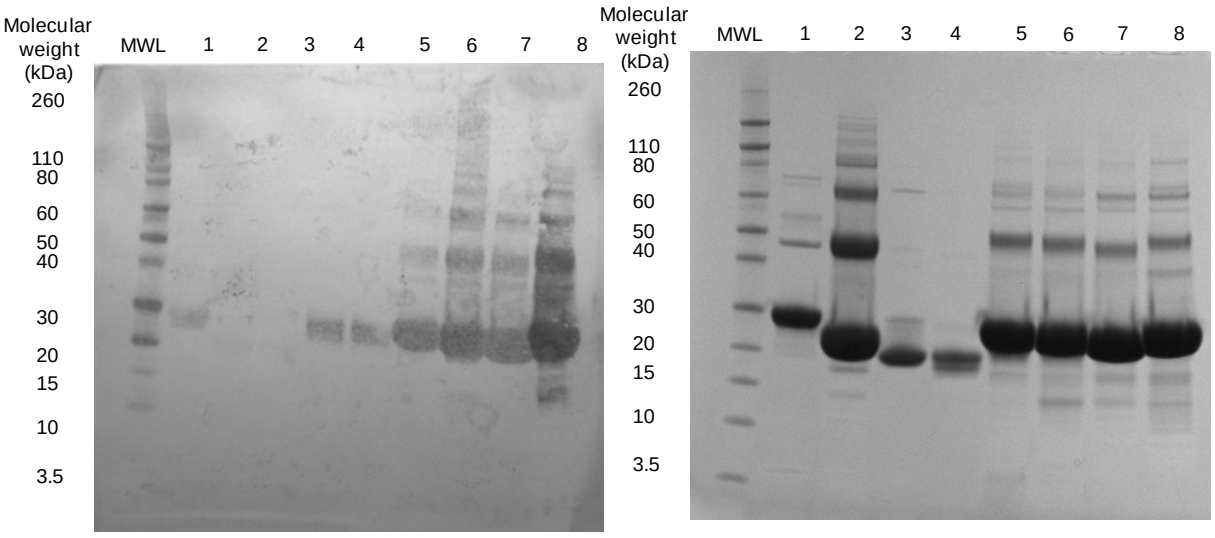


Appendix 5.4: Elution of linker variation constructs from IMAC. (A) PRX3A, (B) PRX3B, (C) PRX3C and (D) PRX3D elution from a Ni^{2+} IMAC column upon increasing imidazole concentration. Black: absorbance at 280 nm (mAU); grey: percentage concentration of 0.5 M imidazole buffer (%). All constructs were expressed and purified successfully. The presence of target protein was confirmed by SDS-PAGE analysis.



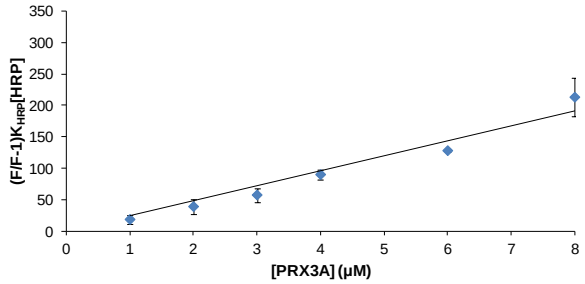
Appendix 5.5: SEC traces of linker variation constructs at 600 mM NaCl. High concentration of salt was added to the protein sample (1 mg/mL) in non-reduced buffer pH 8.0. (A)PRX3A, (B) PRX3B, (C) PRX3C and (D) PRX3D all show a mixture of HMW and LMW oligomers. Increasing the salt concentration from 150 mM to 600 mM does not disrupt the interactions that are driving HMW assembly. Integrated values are shown in table 5.6.

A: western blot **B: SDS-PAGE**

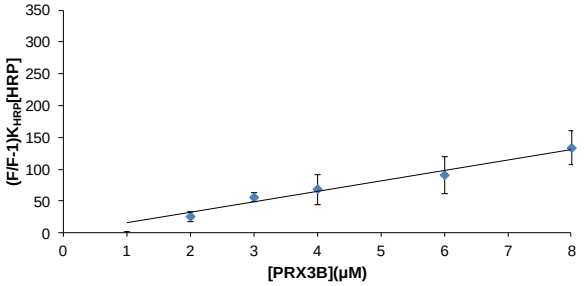


Appendix 5.6: SDS-PAGE (A) and western blot of linker variations. Wells 1 – 4 contain *Hs*Prx3-6his with varying amounts of TEV, related to another experiment, wells 5, 6, 7 and 8 contain PRX3A, PRX3B, PRX3C and PRX3D respectively. There are a small amount of impurities in the first three samples that are not seen on western blot.

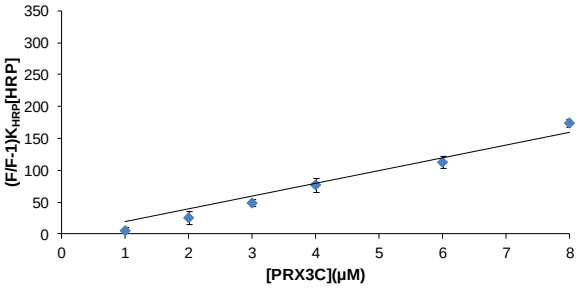
A: PRX3A



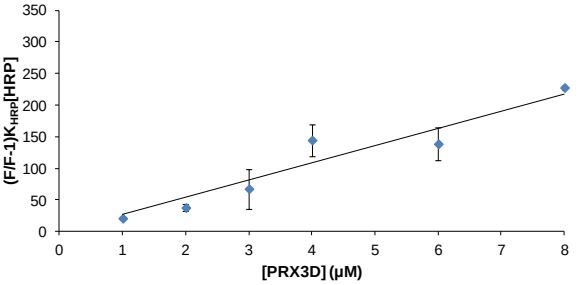
B: PRX3B



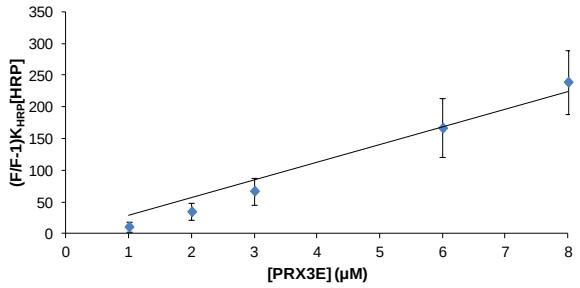
C: PRX3C



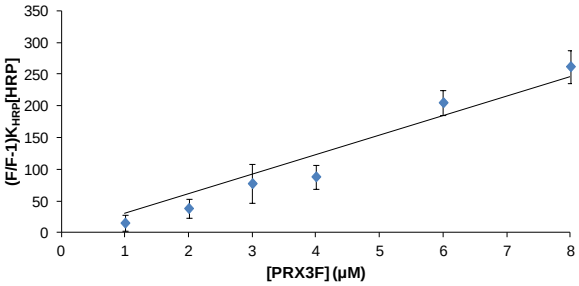
D: PRX3D



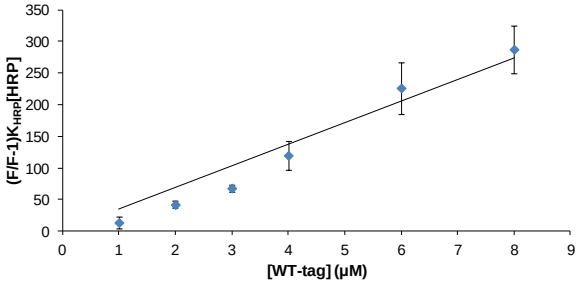
E: PRX3E



F: PRXF



G: *Hs*Prx3-6his



Appendix 5.7: Activity assay analysis. All 7 constructs were analysed using the protocol from the methodology (2.4.1). k_{Prx} was calculated by plotting $\left(\frac{F}{1-F}\right) k_{HRP}[HRP]$ against [Prx], where k_{HRP} = 1.78 M^{-1} s^{-1} [1] and F = the fraction inhibition of HRP.

1. Dolman D, Newell GA, Thurlow MD, Dunford HB: **Kinetic study of reaction of Horseradish-Peroxidase with hydrogen-peroxide**. *Can J Biochem* 1975, **53**(5):495-501.